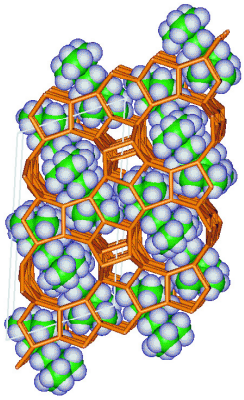
Workflows for Computational Scattering Science 2013



Report from the Community Workshop

January 31 – February 2, 2013
California Institute of Technology

Editors in Chief:

Brent Fultz, John J. Rehr, Simon J.L. Billinge

Topic Editors:

E.E. Alp, S.J.L. Billinge, O. Delaire, B. Fultz, J.Y.Y. Lin,
D. Prendergast, T. Proffen, J.J. Rehr, A.D. Rollett, M. Stalzer, S. Vogl

Disclaimer

This report was prepared as an account of a workshop funded by the National Science Foundation under NSF Award Number 1216716.

Table of Contents

Executive Summary

The spectacular progress in computational methods is well recognized. Less widely known is the spectacular progress in the intensity and brightness of X-rays, neutrons, and electrons that are used for scattering experiments on materials, molecules, and condensed matter. In the past decades, X-ray brilliance has outpaced Moore's Law for computer hardware. These enormous advances in computing and in scattering have occurred independently, and there are now exciting opportunities to combine them to facilitate scientific advances. Computational scattering science aims to do just this, but the full potential of this field has not been realized because appropriate computational tools and resources are not yet in the hands of scattering scientists.

Part of the difficulty in bringing computing to scattering science is the diversity of experiments that are performed with X-rays and neutrons. No single software package can provide advanced data analysis for all users. What is needed to overcome this bottleneck is a set of computational workflows that have proved useful for discovery in scattering science, and a way to adapt these workflows to new types of science. The central engines in these workflows are tools drawn from computational materials science that have found broad use for predicting the structure and dynamics of materials. This report focuses on workflows built around materials simulations based on density functional theory and molecular dynamics. A set of codes beyond density functional theory is emerging, and these need to be developed over the next years, especially for next generation X-ray spectroscopy experiments. Specific use cases are proposed in this report, some as Universal Modeling Language (UML) diagrams, and some as text descriptions. They show paths to transformative science.

Computational techniques facilitate the combined analysis of information from different types of experiments, linked by an underlying model of the structure and dynamics of a material. Such a combined approach requires the assessment of uncertainties in the model and its parameters. For example, it is understood in principle how Bayesian methods can be used to refine prior information as new results are added, and it is understood in principle how to quantify the uncertainties when combining results from similar scattering experiments. Today this is rarely done in scattering science, again because computational tools are not readily available. Furthermore, sometimes the models are not well specified, and new methods of uncertainty quantification are needed from computational and mathematical sciences.

This report also outlines a plan to organize an institute for computational materials science, including staffing and budget. Its scope of science would be a subfield of scattering science. The size and scope of the effort would be matched to that of an efficient team of approximately ten persons. The methods of computational science addressed in this report are primarily those of modern materials physics, specifically methods to calculate the atomic structures of materials, as well as dynamical and excitation properties of their electrons, spins, and atomic vibrations. This covers a wide range of materials research, but by no means all of it. Nevertheless, it is envisioned that several field-specific centers could cover much of the scattering science performed today. It is critically important to establish one such center soon. Proof of a working institute is needed before attempting to serve the entire scattering science community.

1. Introduction

1. Introduction

The past decade has seen extraordinary advances in X-ray and neutron scattering research (see Fig. 1.1). For example, the Spallation Neutron Source (SNS) is operating reliably with an order-of-magnitude increase in neutron flux over its predecessors, and its instruments offer similar gains in efficiency. The Linac Coherent Light Source (LCLS), with brightness six orders of magnitude higher than previous synchrotron sources, has begun operations. The National Synchrotron Light Source 2 (NSLS2) project at Brookhaven National Lab is well underway, and upgrades at the Advanced Photon Source (APS) and the NIST Center for Neutron Research are keeping these facilities competitive for neutron and X-ray scattering science.

Advances in computational methodology are also spectacular. Since 1980, hardware performance per dollar has increased by more than a factor of a million. Software advances have led to a rich field of discovery in computational materials science, and computations that were unimaginable only 20 years ago are commonplace today. There are enormous opportunities for scientific discovery by combining scattering experiments and computational science. The data of Fig. 1.2 show results from a keyword search for publications that combine ab-initio theoretical calculations with scattering experiments. This is, of course, not the whole picture of computing and scattering science because computing is used for every publication from the 14,000 annual users of the synchrotron and neutron facilities in the U.S. Nevertheless, only a few percent of the publications from X-ray

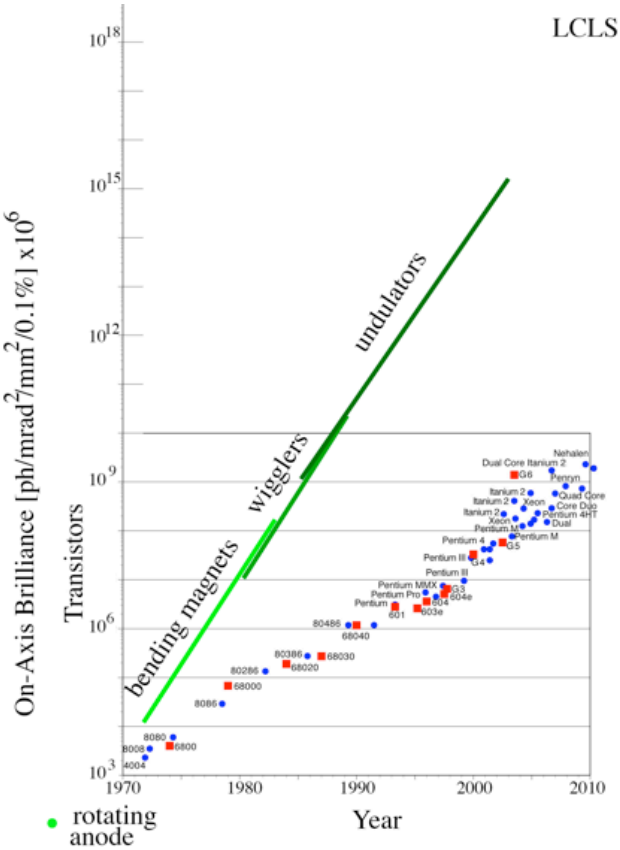


Fig. 1.1. X-ray light source performance versus year, superimposed on a typical plot demonstrating Moore's law (i.e., the exponential increase in the number of transistors on a device).

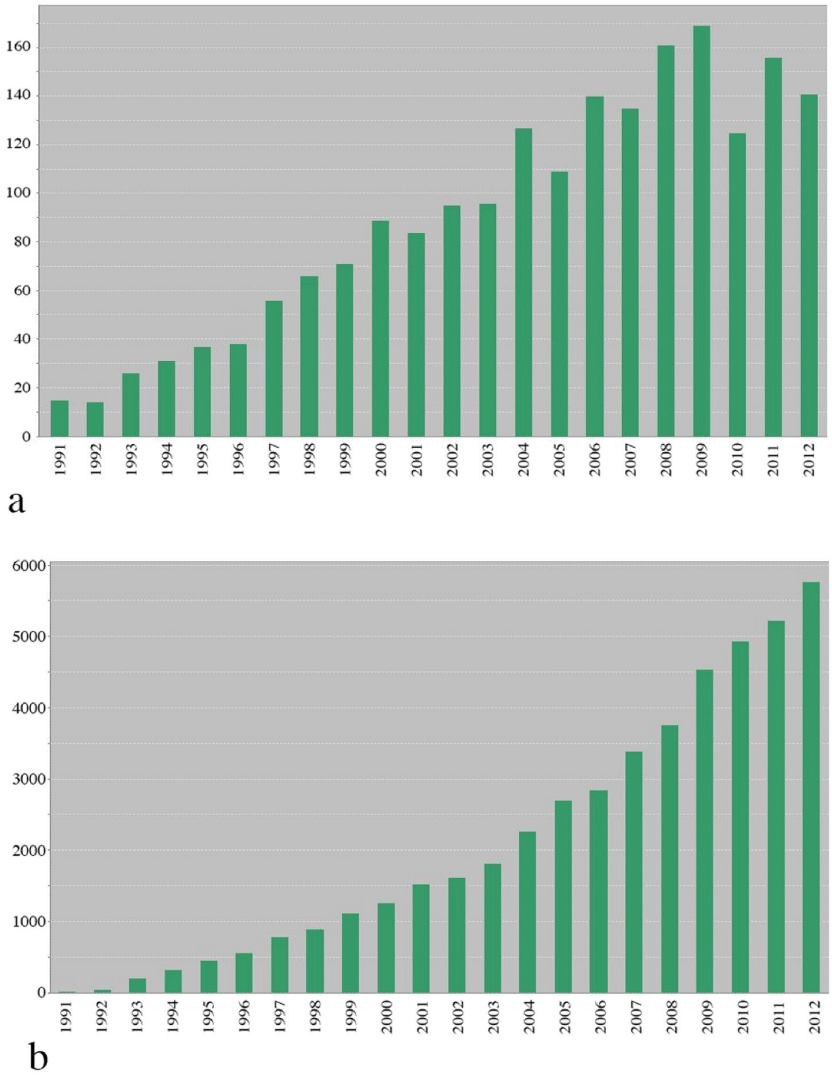


Fig. 1.2. Results of May 2, 2013 from a keyword search of Web of Knowledge, an electronic publications database of Thomson Reuters. The keyword entry was: "ab-initio AND scattering AND (X-ray OR neutron)"
(a) Published papers each year, showing linear trend since 1990. (Total papers is 2018.)
(b) Citations to the papers each year, showing an increasing rate of citation. (Average since 1992 is 23.6 citations/paper)

or neutron scattering research used the more modern tools of computational materials science that were the focus of this keyword search and this report.

The report from a workshop funded by DOE BES and NSF DMR at Argonne National Lab, "Computational Scattering Science 2010", identifies specific fields of science where computation would elevate the understanding of data from scattering experiments [1.1, 1.2]. A related workshop was organized jointly by ASCR and BES of DoE in Bethesda, MD in 2011 [1.3]. The present report is a result of the NSF-supported workshop "Scientific Workflows for Scattering Science," held at Caltech in January 2013 [1.4]. The goal of this recent workshop was to identify specific "computational workflows" that offer best value for discovery in scattering science. The examples include "use cases" of how scattering scientists and computational scientists would work together with modern computational resources. This report identifies natural combinations of computing and scattering measurements that have not been tried, or that exhibit unfulfilled promise.

Today it requires less effort to use sophisticated tools for computational materials science than it did about 10 years ago. Today a graduate student may routinely use more than a million CPU hours per year, and this will increase steadily. Yet there remains a large learning curve for applying these tools to Scattering Science. The bigger challenge is to incorporate these computational tools into workflows for doing new science. It is here where the scattering community needs help to make these tools easier to use. This is the motivation for a Scientific Software Innovation Institute for

X-ray and Neutron Scattering Science (SIXNS). This institute would not focus on developing new codes for computational materials science, but would develop workflows that would largely make use of existing codes for calculating the properties of materials to interpret the results of experimental data by scattering scientists. The institute would build these workflows, adapt them for new scientific endeavors, and support scattering scientists who need them.

It is of course paramount to quantify the accuracy of quantitative science. When sophisticated features are extracted from a workflow that combines computing and experiment, or when different types of experimental data are combined to develop an underlying materials model, we have little experience with the reliability of the results. Uncertainty quantification needs to be considered when building new scientific workflows. Although underutilized in scattering science, Bayesian methods can be incorporated naturally into efforts that combine computation with experiment, using results from one as prior information for the other, for example. These methods are challenged when the models for obtaining conditional probabilities are not well-known, however.

Nearly all scattering science research with X-ray and neutron facilities involves studies of materials, but the experiments cover an enormous range of different materials and phenomena. Nevertheless, some computational materials science methods are important for broad classes of materials, such as methods for calculating electronic structure, molecular dynamics, and tools for modeling atomic structure and dynamics. Software that simulates materials at the atomic level using quantum mechanics and statistical mechanics has become both commonplace and essential for understanding a large range of phenomena. It is time to develop specific scientific workflows that use these modern tools of computational materials science to assist in interpreting specific scattering experiments.

Different topic groups developed the sections of this report, and the group leader was charged with writing them. The editors merged topics into larger sections, but the authors are identified individually. The sections include the following:

- One family of important workflows, described in Section 2, involves simulating experimental data by considering how an undulator or moderator sends X-rays or neutrons into a beamline, how the beams are monochromated and focused, and how they are scattered from the sample into detectors. The quirks of the instrument are included in these simulations, and in many cases it is more reliable to compare simulated data to real data, rather than trying to correct experimental data for characteristics of the instrument. Central to these simulation workflows is the scattering from the sample. We know how a neutron is deflected by different interactions with the atoms in the sample, and we now have the tools to calculate the positions of atoms, their vibrational dynamics, their magnetic moments, and how local magnetic moments respond to temperature and magnetic field. These ab-initio codes are a triumph of quantum mechanics in the past two decades, and they should be used for precise comparison to experiment. An institute would develop the code to calculate correlation functions that give the probabilities of deflections in scattering events, and use these results in the simulations of experimental data.

- Incorporating density functional theory (DFT) codes into workflows is a main topic of Section 3. There is general agreement that DFT codes are the standard workhorses for computational materials science today, and there should be no major issues incorporating them into workflows that calculate diffraction patterns and phonon dispersions of regular crystalline material, for example. To date, many studies have already benefited from DFT workflows. Classical molecular dynamics codes could also be made available this way, but the inputs and outputs from these packages will likely need pre- and post-processing tools that could be provided by an institute. Ab-initio molecular dynamics calculations are becoming increasingly practical; however these methods have seen limited use in computational scattering science largely because of the concern that the results may not justify the substantial computing resources required. Here a computational scattering scientist could help an experimentalist assess the balance between cost and reward, and help initiate such calculations. Finally, there is widespread

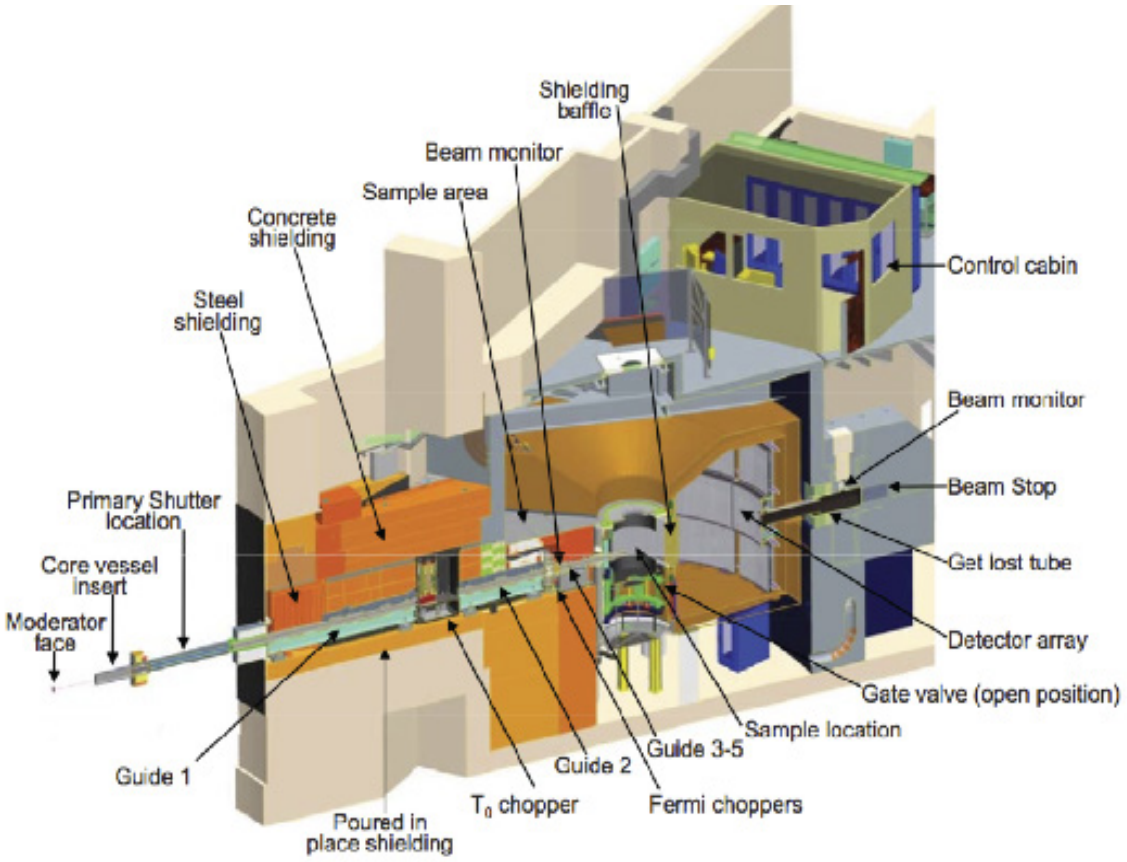
interest, both in theory and in experiment, in understanding materials when their structures are not in equilibrium. Codes to calculate the structures and dynamics of materials with excited electrons are emerging, and are currently limited to much smaller systems than are typical DFT calculations due to the less favorable scaling with system size of these methods. Providing a forum for interaction between theorists and computer scientists would help to establish more efficient implementations, while increasing interactions between theory and experiment would be beneficial to both.

- Structural studies by X-ray diffraction are being performed with much higher spatial resolution than in the past, and a variety of X-ray "microscopes" have recently become available. These instruments can map out strain fields in a material, or map out the three-dimensional microstructure of a polycrystalline material. Section 4 describes how computational materials science codes at the continuum level can be used to interpret the results of such experiments, and how the experiments can further our understanding of microstructural evolution at elevated temperature or in materials under applied stresses.

- Uncertainty quantification is hardly used by the scattering science community, even though it is almost always desirable to estimate errors in experimentally-derived quantities. When these quantities are not obtained by direct measurement, but instead the measurements are used to optimize parameters in an underlying atomic model, for example, it is challenging to estimate the uncertainties. There are some families of workflows where classical Bayesian statistics could be adapted to estimate uncertainty. These methods are not implemented in core packages of computational materials science. The workflows that incorporate these packages need to be designed to include uncertainty quantification. It will be a goal of the institute to develop uncertainty quantification for some standard workflows, especially those that combine multiple sets of data and computation. In other cases the models themselves are uncertain or controversial. These cases may not be analyzed so reliably by Bayesian methods, and Section 5 gives some ideas about how the applied mathematics of uncertainty quantification could help over the next years.

- Section 6 discusses how an institute for computational scattering science could help frame and address the strategic issues that the community faces in hardware and software. The scattering science community can no longer rely on software development by isolated individuals or small groups. This approach dominates in computational scattering science today, as it has for the past three decades. The limitations of this cottage industry approach are now evident, as the field has, with few exceptions, not advanced significantly over the past decades. A community-based institute would need to champion modern methods of code development that allow for archived methods and reusable modules, for example. The community has been well served by advances in microcomputers over the past decades, but this trend is not likely to continue. The needs of consumers and the needs of scientists are not likely to be met by the same hardware, as the consumer marketplace moves its emphasis towards content delivery and storage, whereas scientific hardware moves towards highly parallel architecture to optimize Linpack benchmark performance, for example. The style of scientific programming is likely to change, and an institute can help the community identify the algorithms that will scale well to the new hardware.

  Although the main mission of the software institute would be to incorporate existing software packages into computational workflows, some code development and code modification will always be necessary. Adapting computational tools to new scientific workflows is most efficient when the software packages are modular, and have consistent interfaces for I/O, for example. The rules of good object-oriented programming go a long way towards satisfying this goal, although compatibility must be designed. Also necessary will be workflow documentation, discussion forum management, and some brokerage of computing resources. Software development with professionalism, and with steady input from the scattering community, is essential for building useful and sustainable software for computational scattering science.



The SNS ARCS instrument.

- Section 7 describes how a software institute could implement scientific workflows and make them available to scattering scientists. Every investigation has unique features in its workflow, from data to publication. Expertise in adapting the computational tools in workflows is learned, much like the mastery of experimental tools. A mission of the institute would be to lower the barriers for learning computational tools and their interconnections. In many cases this can occur by scientific collaborations where an experimentalist and a computational scientist work together on a problem with a clear scientific goal. This approach assures that both parties have a shared mission in the workflow, and will share in the scientific discovery. This high-level type of user interaction is considerably different from what is understood as 'user support' in the commercial software industry.

  At the postdoctoral level, many experimental research groups value individuals who can use the tools of modern computational materials science. The SIXNS Institute can offer excellent opportunities for junior computational scientists to do new science by collaborating with experimental groups. There will be a substantial expectation for these junior scientists to publish new science that will advance their careers.

  Another opportunity for workforce development is science education. With good design, such as separating user interfaces from computing engines, software can fulfill the needs of both research and education. With a user interface appropriate for the level of the student, software can help introduce concepts of X-ray, neutron, and electron scattering to undergraduate and graduate students, and to scientists in other fields. Simulations could also show graduate students and scientists if scattering experiments are appropriate for their own research. There is a considerable overlap of requirements for user friendly but advanced software for education, and software for advanced analysis by scattering experts.

The January workshop and this present report go a long way towards showing what could be done by an institute (SIXNS) to elevate the level of scientific computing in X-ray and neutron scattering science, and how the institute would operate. The prior report, Computational

Scattering Science 2010, identified broad scientific opportunities. The present report focuses on specific science thrusts and identifies specific use cases for software workflows that would benefit them. The core function of the institute would be the development of workflows that are needed to do new science in collaborations between experimentalists and computational scientists. The institute would manage the archiving of these workflows, and offer them with modification for new projects as they arise. A national effort would have a broader scope than the single institute proposed here for materials physics; a complete national effort would require institutes that specialize in other fields of science. Nevertheless, in 2013 it is important to get started. There are advantages to initiating a materials physics effort because there is already partial integration of computing into the scattering work today, and the remarkable advances in codes for quantum mechanics of materials offer a fertile field for discovery. Learning how to bring modern computational science to the scattering community needs to be a national priority, and ensuring U.S. leadership requires that the effort begin soon.

[1.1] Workshop on Computational Scattering Science 2010
http://www.its.caltech.edu/~matsci/Publish/CompScatWkshp_2010.html

[1.2] Full Report "Computational Scattering Science 2010"
http://www.its.caltech.edu/~matsci/Publish/CSS_2010.pdf

[1.3] 2011 ASCR/BES Data Workshop
https://www.orau.gov/dataworkshop2011/default1.htm

[1.4] Scientific Workflows for Scattering Science
http://s2i2.caltech.edu/main/

2. Simulations of Experiments

Group Leader: J. Rehr

Group members: E. Alp, J. Beck, K. Broersma, D. Brooks, Y. Cai, R. Car, A. Cunsolo, P. Dai, P. Dawson, O. Delaire, P. Dera, T. Devereaux, P. Duxbury, E. Farhi, R. Ghanem, M. Hagen, M. Hu, M Hybertsen, K. Jorissen, D. Kim, T. Lan, J. Lin, L. Lutterotti, A. Markvardsen, N. Marom, P. Messina, J. Muñoz, R. Neder, R. Osborn, D. Prendergast, D. Reznik, D. Roach, A. Rollett, G. Shipman, J. Tao, B. Toby, J. Tranquada, M. Tucker, L. Udby, S. Vogel, A. Wills and T. Yildirim

2.1. Simulations in Scattering Science

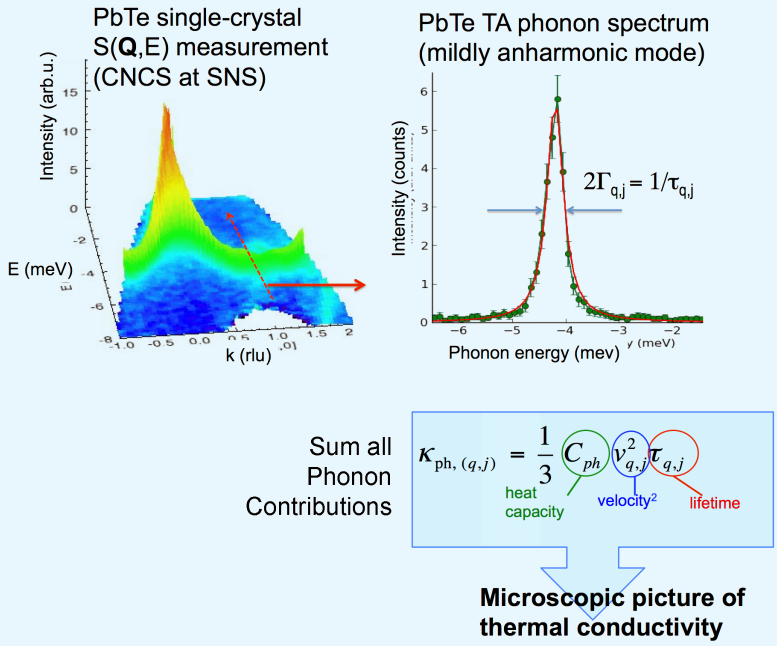
The natural synergy between theory and experiment helps drive scientific developments and is important for research of the highest quality. Computational modeling tools provide a catalyst that fosters this synergy. With the advent of third and fourth generation light sources, neutron scattering facilities, and electron microscopes, experimental probes of unprecedented precision are now possible. Similarly, accurate time-resolved experiments are now possible with XFEL sources like the LCLS. Neutron and electron scattering are important complementary probes of structural and vibrational properties. However, without quantitative theoretical simulations, the capabilities of these modern instruments cannot be fully exploited. Even the best theory is not adequate if the experimental community cannot also apply it. Needed instead is a combination of theory, modeling, and analysis tools that facilitate workflows and speed up the analysis and interpretation of experimental data in all of these scattering sciences.

This Section 2 identifies both topics and workflow scenarios that could eliminate such bottlenecks to scientific progress. The focus is on two prototypical cases that are important for a proposed initiative on computational scattering science encompassing X-ray, electron (ES), and neutron spectroscopies (NS): 1) modeling and validation of XAS and RIXS analysis, and 2) neutron and phonon-spectroscopy simulations and analysis. Use-cases can be developed for ES similar to those for X-rays.

2.2. Use Case: Modeling and Validating for XAS and RIXS Measurements

X-ray Absorption Spectroscopy (XAS) is one of the major techniques for elucidating the structure of materials, and encompasses a number of related techniques. These include EXAFS (Extended X-ray Absorption Fine Structure), XANES (X-ray absorption near edge structure), XMCD (X-ray magnetic circular dichroism), and others. Many of these techniques are now well developed and many theoretical techniques and analysis codes are available. Moreover, new spectroscopies are being developed that take advantage of modern light source technology. In particular, RIXS (resonant inelastic X-ray scattering) is a modern, high-resolution X-ray spectroscopy that takes advantage of tuned X-ray sources to elucidate, properties of materials including both structure, and excited state properties. Photon-in/photon-out spectroscopies such as RIXS yield much higher resolution spectra than is possible with conventional methods, but they require new sources of very high brilliance and more complex beamlines. RIXS theory also requires more sophisticated simulation techniques, since it is based on the Kramers-Heisenberg equation rather than the golden rule, requiring more precise calculations and an explicit treatment of energy losses and the resonant energy denominator. There is now enormous interest in this spectroscopy and demand for analysis tools, which will require substantial extensions of conventional X-ray methodologies. The method is well suited to the goals and capabilities of a software institute.

We propose a use case to develop a Virtual RIXS Beamline, namely, a suite of theoretical and computational tools that can model all of the steps in an experimental RIXS beamline, from the source through the instrumental components and through the sample. As with a real beamline, the tools should interact with each other seamlessly to simulate the experimental results for a given material and measurement characteristics. This virtual beamline concept is at a level of integration higher than that of a typical computational materials science computer code. In this approach, each individual step is represented by a stand-alone code-module, with independent input/output, that can be developed and executed separately and/or in parallel. Rather than enforce any common I/O protocol, we envision the use of data translators that convert from the format of one module to the input of the next. The code modules for a virtual RIXS beamline thus consist of the following:



Computational modeling of neutron scattering phonon measurements. Inelastic neutron scattering provides unique insights into phonon scattering mechanisms and microscopic origins of thermal conduction. (O. Delaire)

1. Model-structures, e.g., from a mix of tabulated databases of experimental data, DFT/MD or more sophisticated ab initio codes, or other model building software. Needed in particular, are the atomic parameters, e.g. xyz, Z, spin, and other physical characteristics such as temperature, pressure, or magnetic field. As with sample preparation, it is important to vet this step against the characteristics of the materials being investigated.

2. Interchangeable electronic structure engines - Modern DFT, DFT/MD, and DMFT codes are needed to produce energy levels, wavefunctions and other data input in spectroscopic calculations. Such DFT engines are widely available today. However, they are typically neither user-friendly nor interchangeable, and making them more so is a major challenge. Needed are interfaces that facilitate their applicability in efficient use-case scenarios. This task is well suited to the framework of our SIXNS initiative and is clearly feasible, as several packages have been successfully integrated as engines into larger scale, automated searching and database generation schemes (some of which contribute to the Materials Genome Initiative). Since these codes often require high performance computational facilities, modest local computer clusters of a few hundred nodes are an important component of such an SIXNS institute. More extensive calculations would be carried out externally.

3. RIXS theoretical engines - Development of efficient and accurate RIXS theories is challenging because the theory requires more sophisticated ingredients, as noted above. Needed are interchangeable theoretical models, including atomic-multiplet, Real-space Greens function (e.g., FEFF), GW/BSE (OCEAN) or other more advanced RIXS theories, as scientists at the Institute or elsewhere develop them. The codes would run on local HPC facilities or possibly high performance or cloud computer platforms. Fortunately a number of the required codes are actively being developed.

4. Data input module - To facilitate data analysis, a module is needed that takes raw experimental data and converts it to physical quantities analogous to $S(\boldsymbol{q},\omega)$. This step also serves to reduce the amount of raw data by preprocessing, and thus is important to reduce big-data bottlenecks. This step alone would be a very useful advance over current approaches.

5. Data analysis modules - Advanced inverse models approaches based on least squares and other a-priori data (e.g., Bayesian and other constraints) to minimize $\chi^2 = |(\chi_{th} - \chi_{ex})^2| + \chi_0^2$ with respect to structural and other parameters. Although there have been a number of efforts along these lines, they

are by no means standard today. They could greatly improve quantitative analysis.

6. Closing the loop - In an optimal workflow scenario, steps 1-5 must be repeated until a specified tolerance is reached. Thus these steps 1-5 have to be automated, and repeated [GOTO 1] until a desirable tolerance is reached.

7. Output module – The output of physical parameters, e.g., structural data, excitation energies, etc. is needed, including visualization tools. This would be part of a user-friendly graphical interface.

For RIXS software developed by SIXNS, a major effort is needed to develop user-friendly interfaces and translator utilities to link these modules. The underlying workflow would first be executed with script-based or command line control. These detailed steps can be concealed in a suitable GUI, based on Java or Web-interfaces, for example. Documentation is essential and must be available online, and a discussion forum may be important, too. Tutorials should be part of the documentation, with complete sample I/O for cases that would be helpful to new users.

2.3. Use Case: Neutron and Phonon-Spectroscopy Simulation

A second illustrative use-case for the scattering science initiative is the simulation of phonon spectra and multi-phonon scattering, allowing for the integration of input from neutron scattering (NS) and theoretical calculations. The software modules in this virtual beam line, much like those for a RIXS beamline, account for all of the components in an experimental NS beamline, all the way from the source to the primary flight path, through the sample, and to the detector, where absolute intensities are recorded and can be compared with real experiments [2.1-2.3]. This can be done either by simulating each neutron trajectory through the instrument, or by breaking the simulation into modules, but the description of the instrument should be as close as possible to reality. This brings in a higher level of science when simulating the sample. The virtual instrument can be operated like a real instrument, and the resulting data analyzed like real data. This concept is implemented, e.g., the VNF portal [2.4] and the Virtual Neutrons for Teaching project [2.5]. (Likewise ray-tracing codes e.g., McXtrace [2.6] exist for X-rays, allowing simulation of instrument performance and using virtual samples from theoretical codes.) Many of the codes needed for simulating samples are the same as those for X-ray instrument simulations, and need not be repeated. A prototypical NS use case would include the following:

- Model-structure software similar to that for the RIXS use case in Sec. 2.2 above.
- Theoretical codes (e.g., advanced DFT codes, MD, etc.) to simulate dynamical matrices and phonon spectra $\omega(\boldsymbol{k})$. Many of the same DFT codes in Sec. 2.2 can be used for this purpose.
- Theoretical codes to simulate, for example, $S(\boldsymbol{q}, \omega)$ or other relevant physical quantities. These would include: a) full simulations of neutron scattering from complex sample assemblies to extract real physical quantities; b) study of contribution of multiple scattering to measured spectra, and methods to remove or account for these artifacts; c) ways to subtract phonon contributions from experimental data to obtain pure magnetic scattering spectra; d) ways to fit simulated spectra to experimental spectra to obtain coupling parameters of a Heisenberg model; and e) uncertainty quantification of the above-mentioned inversion procedure
- Data-manipulation to remove background and produce the experimental signal
- Data-analysis modules

Many of these steps require software development. For example, realizing the various use cases in the third bullet above requires: a) integration of existing tools (mcvine, mcstas, and scattering kernels); b) new scattering kernels, e.g., $S(h,k,l,E)$ for magnetic systems; c) new ways and/or better ways to compute $S(\boldsymbol{Q}, E)$ from first principles; d) new or better ways to compute phonons from first principles; e) integration of an optimizer with the simulation and reduction; and f) integration of UQ framework with optimization, simulation, and reduction.

[2.1] K. Lefmann, et al., J. Neutr. Res. 16 (2010).
[2.2] K. Lefmann, et al., Neutron News 16 (2006).
[2.3] E. Farhi, et al., Collection SFN 12 (2011).
[2.4] http://vnf.caltech.edu
[2.5] http://vnt.nmi3.org
[2.6] http://www.mcxtrace.org

3. Electronic Structure

3.1. Density Functional Theory

Group Leader: O. Delaire

Group members: P. Dai, M. Hu, N. Marom, D. Prendergast, D. Reznik, S. Rosenkranz

3.1.1. Goals

Several scientific areas are fertile grounds for the Institute to develop computation workflows that juxtapose density functional theory (DFT) methods with experimental results. There is a large class of problems where simulation methods could be used to interpret spectroscopic investigations that measure excitations of a system in terms of the underlying structure and chemistry of a material. This theme is expanded below in two use cases focusing on X-ray spectroscopic investigations in connection with electronic structure, and on neutron spectroscopy of phonons as a probe of chemical bonding. Both workflows would also benefit from computational tools that incorporate additional structural information from either experiments or simulations.

It would be most effective for the software effort to implement workflows for use cases that are "low-hanging" fruit, while designing the software in a modular way, creating stepping stones for harder problems later. Central to this strategy, SIXNS should harvest existing DFT/electronic structure codes, and wrap them into reusable modules that can be combined as "Legos". With a focus on providing the "glue" software often missing between key modules (e.g. processing outputs of established DFT codes), a larger scientific breadth can be attained, benefitting a larger community of users. A mode of operation including "collaboration as a service", beyond "software as a service", was also recommended, whereby the expertise of collaborators of the Institute would be leveraged to elevate the modeling skills of the scattering community as a whole.

Another important goal should be to encapsulate "run of the mill" aspects of the computational workflows to facilitate the comparison of first-principles simulations and scattering experiments for a wide spectrum of scientists with either experimental or theoretical backgrounds.

This second aspect raises the challenge of appropriately educating and training users of the workflows to minimize the potential dangers of using DFT codes as black boxes, with uncontrolled approximations. The limitations of DFT methods are well known, but are less obvious to experimentalists. It may seem that lattice parameters, which are so easily measured, can be obtained accurately by the LDA approximation. Similarly, novice users of DFT codes may not recognize that apparent chemical trends of bond strengths in oxides may be useless if electron correlations are important. Thus validating outputs of these DFT modules will be an important task, providing safeguards to the community [3.1].

Bearing in mind the potential pitfalls of turning available DFT codes into black boxes for untrained users, the project should aim to provide unified and simplified interfaces to a range of commonly-used DFT packages. This should promote the modularization of the software, and should enable users with specific needs or preferences to use a particular DFT implementation. Here again, the project should leverage pre-existing developments from the electronic structure community that provide wrappers and tools for pre- and post-processing of calculations. Examples include the Atomic Simulation Environment developed in Europe [3.3], and the ESTEST framework (NSF project) [3.2].

In the above strategy for growing a wider user base for DFT codes in the scattering science community, it will be crucial for the Institute to devote significant resources to training users, and connect theorists with experimentalists. This would include training workshops on computational scattering science and first-principles simulations, and the development of on-line teaching materials. The NanoHub project could provide a good template for an educational platform on computational scattering science [3.2].

3.1.2. Use cases

3.1.2.1. Use case: from X-ray spectroscopy to chemical and structural insights

Theory and experiment complement each other. Spectroscopy is a particularly important case where DFT simulations enable a more complete under-

standing of experimental information(Fig. 3.1). In this use case, structural information is inferred by matching simulated spectra of different structures to the measured signatures. Photoemission and X-ray spectroscopy, combined with theory, can be essential for obtaining structural information. For example, atomic clusters may be too small (or too unstable) to be characterized directly by techniques like X-ray diffraction.

Another example is found in the study of buried interfaces, which are typically characterized only indirectly (if the sample is cleaved, the structure of the buried interface may be destroyed). Simulations of possible structures of a buried interface, and comparison to the measured spectra, can provide insight on the structure and composition of the interface [3.5]. For the computations, the procedure typically includes a global minimum search. In some cases, however, experiments done at elevated temperature or out of equilibrium do not involve atomic structures that are predicted to be most stable by DFT at 0 K [3.4].

3.1.2.2. Use case: inelastic X-ray and neutron scattering studies of phonons in crystalline materials

Spectroscopic investigation of atomic dynamics with inelastic X-ray and neutron scattering is another science area that would benefit from more first- principles modeling support. The thermal, elastic, electronic and chemical properties of materials are intimately related to the atomic vibrations, which bear the signature of forces acting upon the atoms. Accurate determination of phonon frequencies and their dispersions across the Brillouin zone is possible using momentum-resolved inelastic X-ray and neutron scattering with meV resolution. There are currently two instruments at the Advanced Photon Source that are capable of such measurements: HERIX-3 and HERIX-30 spectrometers, as well as about a dozen neutron spectrometers at several neutron scattering facilities in the U.S. Also, by using a nuclear resonance technique, phonon densities of states projected along the incident photon direction can be extracted. This technique, known as nuclear resonant inelastic X-ray scattering, can be performed with sub meV resolution for about a dozen elements including Kr, Fe, Eu, Sn, Dy, Sb, and Te. Inelastic neutron

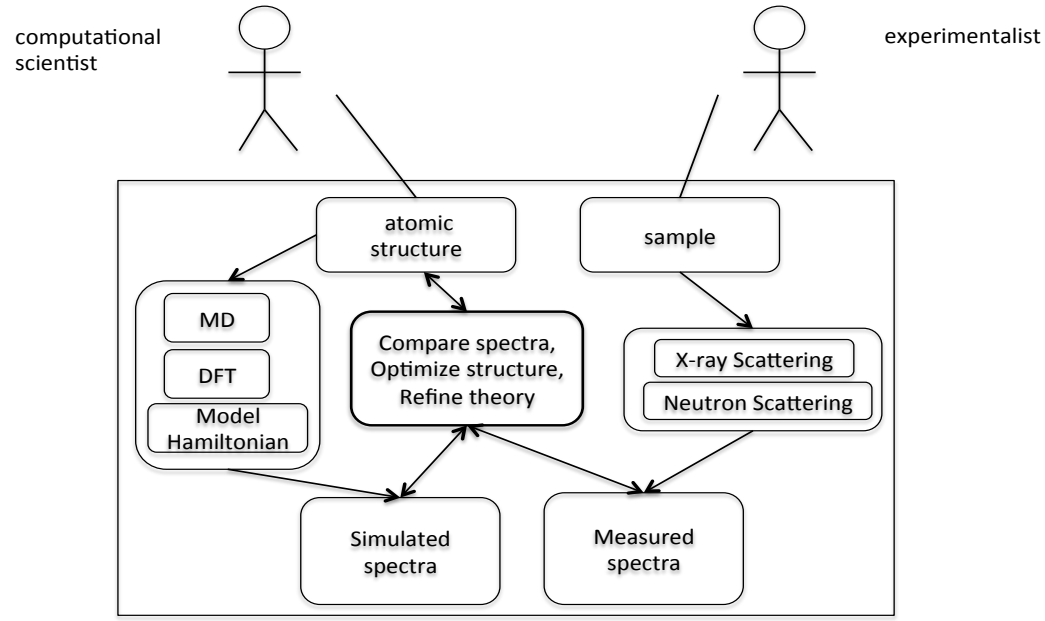


Fig. 3.1. Workflow for modeling experimental spectroscopic data to obtain information about atomic structure.

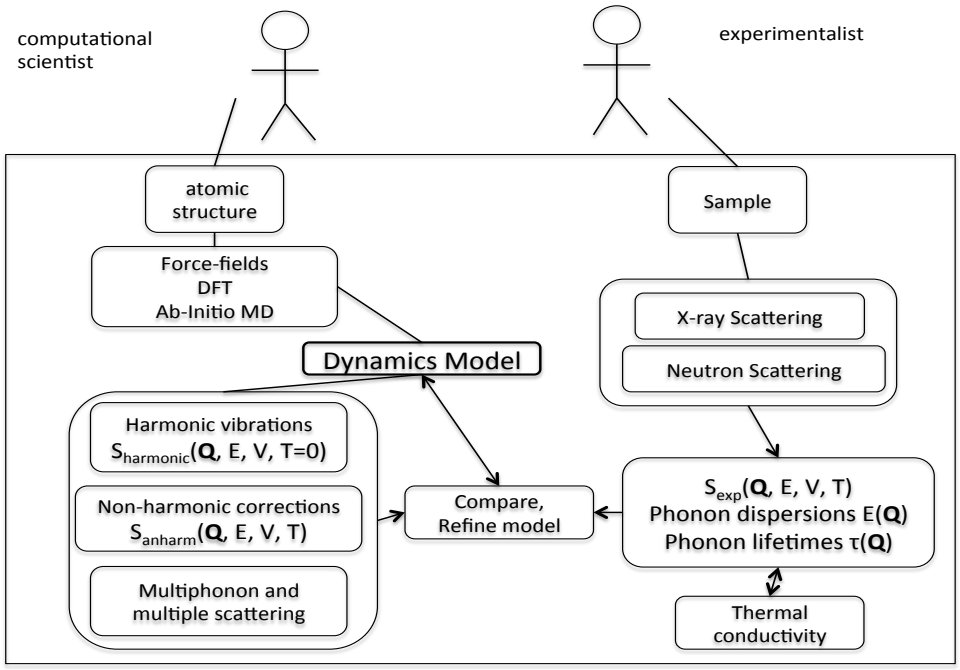


Fig. 3.2. Workflow for modeling of IXS/INS phonon measurements, including anharmonic aspects

scattering measurements are also commonly used to map magnon excitations across reciprocal space, which have most of their intensity at lower momentum transfer, owing to the magnetic form factor. Magnons and phonons often coexist and overlap in reciprocal space, however, and even when an investigation is primarily focused one of the two components, it can be important to take into account the contribution of the other when interpreting the data.

From the phonon measurements with INS and IXS, one can extract the phonon dispersions and their lifetimes, sound velocities and elastic constants, Debye temperatures, Grüneisen constants and the phonon contribution to the free energy. In X-ray experiments, these quantities have been measured under extreme conditions of pressure exceeding 2 Mbar, and temperatures in the range of 5-2500 K. In neutron experiments, they have typically been probed in a wide range of temperatures and under magnetic fields, with some measurements at moderate pressures (below 5 GPa). Flexible computing workflows would allow these results to be interpreted simultaneously with an underlying model, which could be an ab initio computation.

The low-hanging-fruit workflow shown in Fig. 3.2 focuses on deriving reliable phonon and magnon dynamical structure factors for powders and single-crystals. Extensions to the case of coupling between elementary excitations are seen as particularly promising areas where modeling software can enable new scientific insights. First-principles calculations of electronic structure and phonons, when interfaced with measurements of the dynamics, are extremely valuable to relate anomalies in phonon dispersions to the underlying bonding, for example [3.6]. Departures from conventional harmonic lattice dynamics can arise from a rich variety of microscopic couplings accessible with DFT methods, including adiabatic electron-phonon coupling [3.7], phonon- phonon interactions (anharmonicity) [3.8, 3.9], or spin-phonon coupling [3.10]. More fully integrating first-principles simulations with scattering experiments will likely advance our understanding of both the thermodynamics and the transport properties of materials.

3.1.2.3. Collaboration of experimentalists and theorists with the institute

Pertinent to the previous use cases, our group discussed the larger question about the interaction of experimentalists and theorists, and the interaction

of the Institute with national user facilities and their scientists. Productive collaboration channels will require that experimentalists (scattering) and theorists (first-principles) "speak the same language," by making sure they are considering the same physical quantities, and that the experimental and theoretical tools overlap appropriately in time and distance. The proposed Institute should play a central role in facilitating and streamlining this interaction. Specific efforts towards this end are codifying workflows for running software, comparing predictions to benchmarked data, reaching out to a broader theory community, and leveraging expertise of computing experts. A strong emphasis was placed by workshop participants on the important role an Institute would play in coordinating training of the scattering user community by organizing workshops, and strengthening communication between experimental and theory communities. In addition, the proposed institute would need to take on responsibilities in distributing, deploying, and maintaining software to ensure a wide, lasting impact in the scattering community.

3.2. Beyond DFT to Excited Electrons

Group Leader: D. Prendergast

Group members: S. Billinge, R. Car, P. Duxbury, N. Marom, M. McKerns, N. Vast

3.2.1. Overview

Connecting theory directly to experiment, within the context of a Software Institute for X-ray and Neutron Scattering Science, naturally requires accurate descriptions of excited electronic states once we consider resonant or inelastic processes. In addition, monitoring the time evolution of particular non-stationary or non-equilibrium electronic excited states of relevance to energy transfer or conversion processes, even with elastic scattering, requires a fundamental understanding of electrons beyond their ground state. Density functional theory (DFT), described in the previous section, is a ground state theory. DFT exhibits known limitations related to local or semilocal approximations, derivative discontinuities, and missing dispersion forces, for example. Additionally, it may be difficult to extend the effective models of DFT such as Kohn-Sham orbitals to experimentally-probed excitations. Excited state computational methods exist and are being applied in various contexts, but are not yet at the level where a casual user can easily and confidently apply them within a computational workflow.

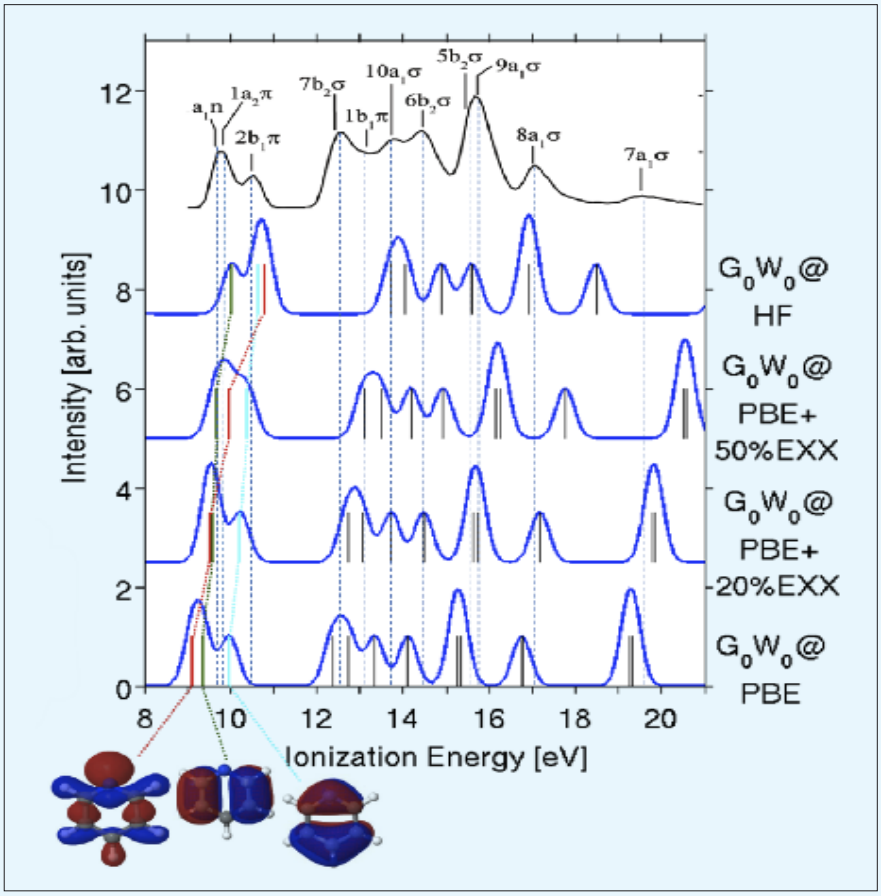
Today the excited state methods described in this section (from the many-body physics community, as we will not discuss quantum chemistry methods) have not yet reached the level of black box implementation of DFT. The theoretical equations and formalism have not changed much during that past three decades, yet we still lack consensus on the accuracy and consistency of results generated by excited state methods, due to subtle specifics related to numerical representation, convergence parameters, and physical approximations. This is exacerbated by the difficulty of the experiments themselves – spectroscopy is notoriously sensitive to factors that may be beyond control. The excited state methods lack the advantage of DFT for direct comparison to X-ray diffraction bond lengths and thermochemistry for formation energies. Therefore, it is perhaps prudent to establish within the Institute a body of expertise to educate the scientific community at large on excited electron methods and associated algorithms, and work with computer scientists and applied mathematicians to overcome computational bottlenecks. Also appropriate would be a forum for open and evolving discussion and execution of appropriate benchmark calculations to build a repository of acceptably well-characterized systems.

The electronic excitations considered here are limited to singly charged and neutral excitations. Singly charged excitations refer to electron addition or removal energies and their associated amplitudes. Within the Hartree-Fock approximation, the individual molecular orbital eigen-energies correspond to these same excitations, via Koopman's Theorem. However, DFT has no such theorem associated with its Kohn-Sham eigen-energies, and only the highest occupied orbital energy is guaranteed to approximate the true ionization potential (IP). Instead, many-body perturbation theory employs the GW approximation to model such excitations, as described below. In this case, the single electron excitations and their associated polarization responses are called quasiparticle (QP) excitations. The spectrum of QP excitation energies corresponds directly to the electronic band structure and is measurable via several techniques: valence band (or ultraviolet) photoemission spectroscopy for the occupied states close to the Fermi level, inverse photoemission spectroscopy for the unoccupied states close to the Fermi level, and X-ray photoelectron spectroscopy for deep inner-shells or core-levels.

Only neutral excitations that can be described within an effective two-particle model based on correlated electron-hole QP pairs are considered in this summary. These are the states that might be accessible to optical absorption spectroscopy in a wide energy range (from IR to X-rays) that specifically excites electrons (and not phonons, magnons, etc.). A time-dependent reformulation of DFT, developed by Runge and Gross, provides one solution to this problem with limitations based on our current lack of accurate exchange-correlation kernels. Additionally, the Bethe-Salpeter equation (BSE) offers a natural extension of the GW QP approximation to describe coupled two-particle excitations, whose mutual electrostatic attraction can lead to the formation of bound excitonic states.

3.2.2. Many-body perturbation theory: GW and BSE

Many-body perturbation theory may provide a reliable description of scattering experiments



A consistent starting point for G0W0. Figure shows spectra of pyridine based on different DFT functionals, compared to PES. (N. Marom)

that cause electronic excitations, such as UV and X-ray photoelectron spectroscopy (PES, XPS), UV-visible absorption spectroscopy, and X-ray absorption spectroscopy (XAS) [3.11]. The GW approximation, where G is the one-particle Green's function and W is the dynamically-screened Coulomb interaction, is the current state-of-the-art for simulating electron emission processes (PES, XPS) [e.g., see Fig. 3.3.] GW accounts for the electronic response to the emission of an electron, which is absent from ground state density functional theory (DFT); as such, GW provides a highly accurate description of band gaps [3.12, 3.13], defect/ dopant charge transition levels [3.14], and the energy level alignment at interfaces [3.15-3.20]. To describe photon absorption processes, one must also account for electron-hole interactions. This may be achieved by using the output of a GW calculation to construct the two-particle Green's function, and then solving the Bethe-Salpeter equation (BSE) [3.21]. BSE is equivalent to time dependent DFT (TDDFT) in the electronic process it describes. BSE has been shown to be more reliable than TDDFT for describing charge transfer excitations, and is considered the state-of-the- art for simulating photon absorption processes.

Performing GW/BSE calculations requires a high level of technical skill and understanding of the physical and numerical approximations involved. These approximations may be sources for errors, and must be carefully tested for each system on a case-by-case basis. GW calculations may be performed at different levels of self-consistency [3.22]. The non-self-consistent G0W0 scheme [3.12] is most often used because it provides good

accuracy at a lower computational cost. The partially self- consistent quasiparticle self-consistent GW (QP- scGW) scheme [3.23] is also becoming increasingly popular. However, the results of non-self-consistent and partially self-consistent GW schemes depend strongly on the quality of the underlying DFT calculation [3.22]. Starting a GW calculation from an inappropriate DFT method may cause significant qualitative errors [3.22, 3.24-3.28]. Other sources of errors in GW calculations are basis set convergence[3.29, 3.30] and the various approximations used for calculating the self-energy (e.g., plasmon pole models, analytical continuation, and contour deformation) [3.31]. Errors stemming from these sources typically reduce the quantitative accuracy of the affected calculations without severe qualitative consequences. In plane-wave implementations of DFT and GW/ BSE the core electrons are replaced by pseudo- potentials. Pseudo-potentials may cause small errors for valence states [3.32-3.34]. However, for accurate simulations of X-ray spectroscopy experiments, core level excitations must be treated explicitly. This requires using all-electron implementations without pseudo-potentials. To avoid the above-described pitfalls, appropriate training must be provided for users who wish to embark on GW/BSE calculations.

3.2.3. TDDFT

The ability to compute the Energy Loss Spectra (EELS) of valence electrons from first principles is of fundamental importance, both to complement and help interpret experiments, and to predict the plasmon energies of new materials. A commonly-used approach to compute plasmons is Time-Dependent Density Functional Theory (TDDFT) [3.35, 3.36]. Plasmons at the surface of metals play a crucial role in physics, materials science, and biology. Excitation of a surface wave by a laser on a carefully-designed-and- nanostructured surface is a promising mechanism for the generation of a proton beam, useful in medical applications that require the production of proton isotopes like positron emission tomography or proton oncology [3.36]. In physics, plasmonics is photonics based on collective electronic excitations localized at the surface of the material, which should enable a significant reduction in size of optoelectronic devices down to the nanometer scale. Plasmonics has recently been highlighted in mainstream media through the concept of a "cap of invisibility", as the use of carefully designed nanostructures enables an optical camouflage that makes an object invisible in a given frequency range [3.37]. Last but not least, "acoustic" surface plasmons, whose frequency goes to zero when the wavelength increases, have been recently observed [3.38]. They offer the possibility of light confinement at locations on the surface, with possible applications in photonics and nanooptics [3.39]. Thus, it is highly desirable to develop theoretical methods and computational techniques to obtain EELS spectra that are both accurate and scalable to systems with a large number of atoms, like surfaces, and to bridge the gap between methods accurate for valence EELS and those useful for the spectroscopy of core electrons, like ELNES. One step in this direction has been made recently. A new method based on the TDDFT has been developed to compute the electronic response and evaluate the collective oscillation frequencies at various wavelengths [3.40, 3.41], up to 100 eV, opening the door for ab-initio calculations of plasmons in large systems without any fitting parameter.

3.2.4. Use case: excited state expertise

We imagine creating a web-based forum (Fig. 3.3) where experts in excited state theory and computation can post their codes (or links to external sites) and details of their methodology, publish results of benchmark calculations with associated necessary metadata to repeat the test, and query or validate the best experiments of relevance to these excited state calculations (such as X-ray spectra). This would require the necessary web infrastructure and development to support all this information and also high performance computing (HPC) resources to perform the computational tests. Non-experts who might benefit from or contribute to such a forum include:

• Computer science or mathematics experts who can look directly for algorithmic details of the method and check scaling in cost within the benchmark database to determine how the methods might be improved.

• DFT experts who develop or apply the base codes that provide Kohn-Sham eigen-value and eigen-state inputs for excited state methods.

• Experimentalists who have relevant data they would like to share with potential collaborators for much needed interpretation, who are merely curious about what can be done in theory to describe excited states, and who may be interested in collaborating on future projects.

The responsibilities of the excited electron expert would include collection of various benchmark/experimental data from the literature, and convincing external scientists and their groups to contribute to the Forum as honorary guest members. Other tasks would be regenerating results with available HPC resources, and engaging the CS community to design improvements in efficiency or scaling of current implementations.

3.2.5. Use case: connection of excited state methods to experiment

This use case covers the specific connection (Fig. 3.4) between excited state calculations and useful output for comparison with experiment. It is analogous to the discussion at the end of subsection 3.1 on DFT collaborations on methods and workflows. What would differ would be the level of involvement of institute staff in actually performing the nontrivial calculations, or the development of collaborative interactions with external groups who are recognized experts in the field of first-principles excited-state computational methods. It would involve consultation with experimental groups to define the experiment, associated experimental uncertainty (e.g. resolution), plotting format, etc. It would also require development by the excited electron expert to add carefully-derived broadening of spectral features based on structural disorder, conformational dynamics, electron-phonon coupling constants, etc. Finally, specific experiments to be modeled would include photoemission for all GW calculations, and angle-resolved photoemission, which can reveal band structure and dispersion. Note that solutions at all electron wave vectors $\boldsymbol{k}$ are typically not accessible because the non-local GW Dyson equation provides solutions on a grid, based on accumulation of information on the same grid. The development of intelligent Brillouin zone interpolation schemes would accelerate such methods.

3.3 Fast Dynamics

Group leader: E.E. Alp

Group members: R. Car, A. Cunsolo, P. Dera, T. Devereaux, P. Duxbury, R. Ghanem, D. Kim, L. Mauger, D. Prendergast, J. Tao, N. Vast, L. Young

3.3.1. Emerging experimental capabilities

Most experiments today are conducted in the regime where electron and ion dynamics are separated, but this is changing with the arrival of new picosecond to femtosecond intense probes of X-rays and swift ion beams. These bring new challenges for theoretical predictions of structure at the atomistic and electronic level. It will be necessary to extend materials simulations to accommodate extreme conditions of temperature (~10^6 K) and extremely short times (10^{-12} to 10^{-15} s).

There is a major international investment in facilities to do science on short timescales, and a number of facilities have recently become operational for studies of fast dynamics. These include:

• FLASH, FLASH-II: Free Electron Laser emitting light between 30 and 300 eV, located in Hamburg, DESY.
• LCLS, LCLS-II: Linear Coherent Light Source, operating in the energy range of 8 keV
• SACLA: Japan, SPring-8 Angstrom Compact free electron laser, emitting light at 4-27 keV range, also providing an interaction point between a storage- ring produced X-ray and FEL.
• FERMI: Free electron laser for multidisciplinary investigations, emitting light in the range of 12-124 eV

Several other facilities will be commissioned soon, including:

• EuroXFEL: European X-ray free electron laser in Hamburg, emitting light at 200 eV- 24 keV (expected to be operational by 2016)
• SwissFEL: Paul Scherrer Institute, Zürich, Switzerland, emitting light between 120 eV – 8 keV, (expected to be operational by 2016)
• PAL-XFEL: Pohang free electron laser facility, emitting light at 8 keV, (expected to be operational by 2014)
• JPARC: Japan proton accelerator research complex, Ibaraki,
• FRIB: Facility for Rare Isotope Beams, East Lansing, Michigan

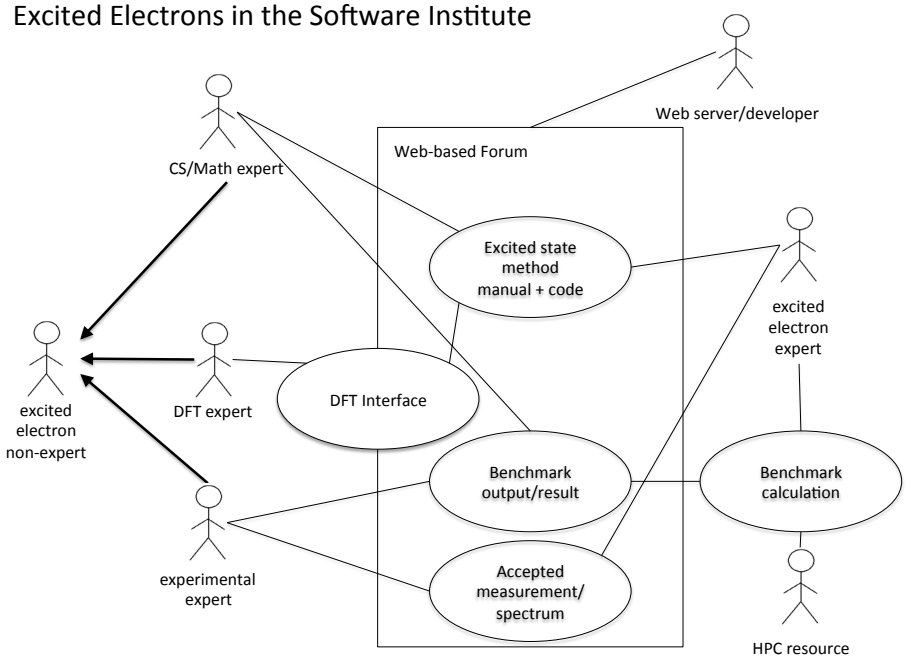


Fig. 3.3. Use Case for the Excited Electron Forum of a Software Institute, indicating all relevant actors, but focusing on the excited electron expert class as key to operation of this aspect of an Institute.

3.3.2. Use case: swift heavy ion-matter, and intense photon field-matter interactions

It is expected that behavior of materials during the interaction of energetic ions can create extreme conditions with deposited energy at the level of 10,000 eV/nm, and 10-100 eV/atom. Such heavy loading may create conditions that can be best described by the so-called "two-temperature model" of electron relaxation in non-equilibrium thermodynamics, where electron and phonon dynamics are decoupled upon excitation. On short time scales, the slow process of electron-phonon coupling is arrested, and can be viewed with different probes. Our theoretical models need to be improved to match these new experimental capabilities. Such models can later be used to explain material behavior under shock waves, similar to conditions created at facilities like NIF. This work will also have implications in developing radiation resistant detectors, materials for space missions, and other strategic applications.

The adaptation of DFT methods to femtosecond time scales at high temperatures will be transformative. Of particular interest is establishing the transition region between slow and fast dynamics. Molecular simulation is a powerful tool for exploring such atomistic processes, but current methods neglect essential effects of energetic electrons. These effects can be accounted for in molecular dynamics (MD) simulations by coupling the atomic motion to a continuum representation of the electrons through a two-temperature model, where the electronic subsystem is represented with the heat equation solved on a grid to mimic the thermal response of the excited carrier electrons. The two subsystems are coupled through energy exchanges prescribed in the two-temperature model.

3.3.3. Use case: thermalization of photoelectrons and squeezed phonons, presence of charge density waves

Squeezed states are quantum states of a harmonic oscillator that exist when the variances of two conjugate variables oscillate out of phase with each other. Squeezed phonon states can be created with ultrafast excitations by a laser pulse with duration much shorter than a phonon period, causing variance of the atomic displacements due to a sudden change in the interatomic bonding strength. Lattice dynamics in out-of-equilibrium conditions generated by fast probes pose a challenge for DFT codes, as the electronic distribution is unknown at these short times. Fast dynamics are at the root of many modern materials issues confronting materials scientists and condensed matter physicists. The dynamics of charge density waves in photo-excited materials, and the interplay between long range-charge order, magnetism and superconductivity are only a few examples.

3.3.4. Use case: chemical and solvation dynamics in picosecond regime

With the prospect of tunable, high repetition rate picosecond synchrotron radiation sources at LCLS-II and SACLA, FLASH-II, and soon at EuroXFEL, new areas of research will come to the forefront. In addition to new sources in the X-ray wavelength region, new and powerful sort-pulse visible light sources have emerged. A billion dollar European project named "Extreme Light Infrastructure", ELI, will include an attosecond short powerful laser in Szeged, Hungary. This ELI pillar will be dedicated to extremely fast dynamics by taking snap-shots in the attosecond scale of the electron dynamics in atoms, molecules, plasmas and solids.

We can anticipate the following types of experiments:

- Visualization of unoccupied and occupied electron orbitals using time-resolved X-ray absorption and X-ray emission spectroscopy,
- Combined time-resolved X-ray absorption spectroscopy,
- Stimulated X-ray emission spectroscopy, and
- Resonant inelastic X-ray spectroscopy studies on chemically-reactive systems.

It is premature to specify software requirements for scientific interpretations with these methods, since they are still emerging. Nevertheless, it is clear that scientific discovery with these new techniques will require computational methods beyond simple reduction and visualization of experimental spectra.

[3.2] https://nanohub.org/
[3.3] https://wiki.fysik.dtu.dk/ase/
[3.4] Noa Marom, Minjung Kim, and James R. Chelikowsky, Phys. Rev. Lett. 108, 106801 (2012).
[3.5] B. D. Schultz, N. Marom, D. Naveh, X. Lou, C. Adelmann, J. Strand, P. A. Crowell, L. Kronik, and C. J. Palmstrøm, Phys. Rev. B. 80, 201309(R) (2009).
[3.6] F. Weber, S. Rosenkranz, J.-P. Castellan, R. Osborn, R. Hott, R. Heid, K.-P. Bohnen, T. Egami, A. H. Said, and D. Reznik, Phys. Rev. Lett. 107, 107403 (2011).
[3.7] Olivier Delaire, Karol Marty, Matthew B. Stone, Paul R. C. Kent, Matthew S. Lucas, Douglas L. Abernathy, David Mandrus, and Brian C. Sales, Proc. Nat. Acad. Sci. USA 108, 4725-4730 (2011).
[3.8] Chen W. Li, Xiaoli Tang, J. A. Muñoz, J. B. Keith, S. J. Tracy, D. L. Abernathy, and B. Fultz, Phys. Rev. Lett. 107, 195504 (2011).
[3.9] O. Delaire, J. Ma, K. Marty, A. F. May, M. A. McGuire, M-H. Du, D. J. Singh, A. Podlesnyak, G. Ehlers, M. D. Lumsden, and B. C. Sales, Nature Materials 10, 614-619 (2011).
[3.10] T. Yildirim, Phys. Rev. Lett. 102, 037003 (2009).
[3.11] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. 74, 601 (2002).
[3.12] M. S. Hybertsen and S. G. Louie, Phys. Rev. B 34, 5390 (1986).
[3.13] M. Jain, J. R. Chelikowsky, and S. G. Louie, Phys. Rev. Lett. 107, 216806 (2011).
[3.14] M. Jain, J. R. Chelikowsky, and S. G. Louie, Phys. Rev. Lett. 107, 216803 (2011).
[3.15] I. Tamblyn, P. Darancet, S. Y. Quek, S. A. Bonev, and J. B. Neaton, Phys. Rev. B 84, 201402 (2011).
[3.16] J. B. Neaton, M. S. Hybertsen, and S. G. Louie, Phys. Rev. Lett. 97, 216405 (2006).
[3.17] J. M. Garcia-Lastra, C. Rostgaard, A. Rubio, and K. S. Thygesen, Phys. Rev. B 80, 245427 (2009).
[3.18] J. M. Garcia-Lastra, C. Rostgaard, A. Rubio, and K. S. Thygesen, Phys. Rev. B 81, 049901 (2010).
[3.19] K. S. Thygesen and A. Rubio, Phys. Rev. Lett. 102, 046802 (2009).
[3.20] C. Freysoldt, P. Rinke, and M. Scheffler, Phys. Rev. Lett. 103, 056803 (2009).
[3.21] M. Rohlfing and S. G. Louie, Phys. Rev. B 62, 4927 (2000).
[3.22] N. Marom, F. Caruso, X. Ren, O. T. Hofmann, T. Körzdörfer, J. R. Chelikowsky, A. Rubio, M. Scheffler, and P. Rinke, Phys. Rev. B 86, 245127 (2012).
[3.23] M. van Schilfgaarde, T. Kotani, and S. Faleev, Phys. Rev. Lett. 96, 226402 (2006).
[3.24] T. Körzdörfer and N. Marom, Phys. Rev. B 86, 041110(R) (2012).
[3.25] N. Marom, X. G. Ren, J. E. Moussa, J. R. Chelikowsky, and L. Kronik, Phys. Rev. B 84, 195143 (2011).
[3.26] E. Salomon, P. Amsalem, N. Marom, M. Vondracek, L. Kronik, N. Koch, and T. Angot, Phys. Rev. B 87, 075407 (2013).
[3.27] L. Y. Isseroff and E. A. Carter, Phys. Rev. B 85, 235142 (2012).
[3.28] P. Liao and E. A. Carter, Phys. Chem. Chem. Phys. 13, 15189 (2011).
[3.29] B. C. Shih, Y. Xue, P. H. Zhang, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. 105, 146401 (2010).
[3.30] S. Sharifzadeh, I. Tamblyn, P. Doak, P. T. Darancet, and J. B. Neaton, Eur. Phys. J. B 85, 323 (2012).
[3.31] M. Stankovski, M. Stankovski, G. Antonius, D. Waroquiers, A. Miglio, H. Dixit, K. Sankaran, M. Giantomassi, X. Gonze, M. Côté, and G.-M. Rignanese, Phys. Rev. B 84, 241201 (2011).
[3.32] P. Umari and S. Fabris, The Journal of Chemical Physics 136, 174310 (2012).
[3.33] R. Gomez-Abal, X. Z. Li, M. Scheffler, and C. Ambrosch-Draxl, Phys. Rev. Lett. 101, 106404 (2008).
[3.34] E. Luppi, H. C. Weissker, S. Bottaro, F. Sottile, V. Veniard, L. Reining, and G. Onida, Phys. Rev. B 78, 245124 (2008).
[3.35] E. Runge and E. K. U. Gross, Phys. Rev. Lett. 52, 997 (1984).
[3.36] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. 74, 601 (2002).
[3.37] Teresa Bartal, Mark E. Foord, Claudio Bellei, Michael H. Key, Kirk A. Flippo, Sandrine A. Gaillard, Dustin T. Offermann, Pravesh K. Patel, Leonard C. Jarrott, Drew P. Higginson, Markus Roth, Anke Otten, Dominik Kraus, Richard B. Stephens, Harry S. McLean, Emilio M. Giraldez, Mingsheng S. Wei, Donald C. Gautier and Farhat N. Beg, Nature Physics 8, 139 (2012).
[3.38] Y. Wang, E.W. Plummer and K. Kempa, Advances in Physics 60, 799 (2011).
[3.39] B. Diaconescu, K. Pohl, L. Vattuone, L. Savio, P. Hofmann, V.M. Silkin, J.M. Pitarke, E.V. Chulkov, P.M. Echenique, D. Farias and M. Rocca, Nature 448 (2008) 57.

[3.40] I. Timrov, Ab initio study of plasmons and electron-phonon coupling in bismuth : from free-carrier absorption towards a new method for electron energy-loss spectroscopy, PhD manuscript in preparation.
[3.41] Iurii Timrov, Nathalie Vast, Ralph Gebauer, and Stefano Baroni, Liouville-Lanczos approach for electron energy-loss spectroscopy within linear-response time-dependent density-functional perturbation theory, to be submitted (2013).
[3.42] M. Beye, S. Schreck, F. Sorgenfrei, C. Trabant, N. Pontius, C. Schüßler-Langeheine, W. Wurth and A. Föhlisch, Stimulated X-ray emission for materials science, Nature 501, 191 (2013).

4. Nanostructure and Microstructure

Group Leader: A.D. (Tony) Rollett,

Group Members: P. Dawson, P. Duxbury, L. Lutterotti, R. Neder, T. Proffen, S. Rosenkranz, B. Suter, J. Tao

4.1. Overview of the Field

Improved computational workflows will lead to new understanding of how the nanostructure and microstructure of materials determine material properties. For example, lifetime reliability of structural components such as batteries, fuel cells, vehicles and so on, are limited by fatigue, fracture, and creep failure, in complex ways. This section describes how newly-developed experimental scattering techniques are in need of advanced simulations of microstructures and their dynamics to elevate our understanding of microstructure-property relationships.

Understanding how microstructure affects properties requires information on evolving 3D structures. Today, however, much is based on static pictures and 2D data, even though the full 3D microstructure is essential for predictions. There is an emerging effort to bridge the time and length scales of materials by linking density functional theory and molecular dynamics calculations to the nanoscale, linking the nanoscale to a continuum, and linking dislocation dynamics to continuum methods in finite element calculations, for example. These simulation efforts are large in their own right, but as they emerge it is important to have tools and methods to connect their predictions to experimental measurement with enough detail to address individual steps in the models.

There are important connections between these computations with scattering measurements and in situ testing of mechanical response. Examples include

- Creep experiments where linking damage accumulation directly to microstructure can provide more quantitative constitutive models for components lifetimes.
- Investigation of superplastic deformation to partition between mechanisms such as grain boundary sliding and dislocation creep.
- Investigation of the distribution of martensite variants in shape memory alloys.
- Twinning frequency and distribution in advanced steels.

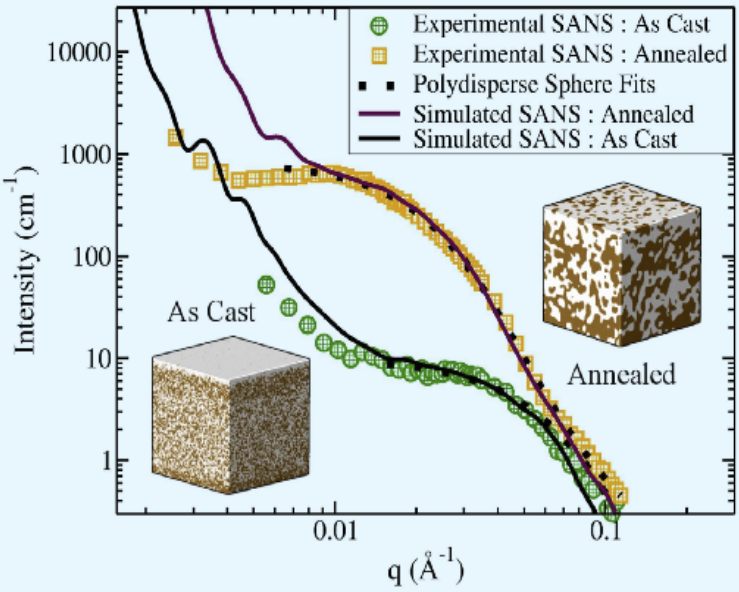
If point-wise mapping of elastic strain (as is presently possible at the scale of grains [4.1]) could be added to the analysis, yet more applications would open up, simply because so many engineered materials are polycrystalline and have complicated microstructure-property relationships that are affected by their 3D structures. In this field of engineering materials research, combinations of techniques have proved most effective for problem solving. Combining energy analysis with tomography or diffraction or imaging is likely to be increasingly important to researchers.

Another service that could be provided by a computational scattering science institute is the expertise to allow researchers to explore methods for applying scattering and simulation methods to their problems. Working with computational scientists and scattering scientists will prompt users to consider alternate experimental methods or combinations of methods (e.g. tomography with diffraction) that they might not think of for themselves. Some planning services can be offered by software workflows, such as a system that provides virtual instruments and virtual training on them. (An example is the Virtual RIXS beamline of Section 2.) Such an assessment of scattering methods may help with a cost/benefit analysis of different approaches to the problem. Today the barrier is high for adding uncertainty quantification to an analysis, and an institute could persuade and direct users to add UQ for increased confidence in the work. Finally, over time there will be an archive of workflows and microstructural data. In principle, this archive could offer explorations of its own.

4.2. Use Cases

4.2.1. Demonstration of HEDM virtual experiment

High energy diffraction microscopy (HEDM) [4.2] involves data acquisition on the terabyte scale with a 1024x1024 tiff image at two distances, multiplied by 180 angular positions, multiplied by approximately 100 layers of a sample of 1 mm^2 cross-sectional area. This is preceded by an alignment procedure and followed by a computationally intensive analysis that performs segmentation on the images, with forward modeling (simulated annealing)



Tools for computational scattering science. (P. Duxbury)

to reconstruct the orientation field in each layer [4.2]. Prior to the actual reconstruction calculation, which dominates the consumption of cycles, an optimization must be performed on the experimental parameters (e.g. specimen-to-detector distance). Turning the alignment procedure into a virtual experiment so that new users could practice before their beamtime would be useful. Guided simulations of how to perform the parameter optimization and image segmentation would be extremely useful to both potential users and occasional practitioners. The use case in this example would involve users, the software, and beam scientists to answer questions about the alignment phase and expert users to answer questions in the data analysis phase.

Access to 3D orientation maps measured non-destructively by HEDM opens up many possible applications, especially if combined with computer tomography. Examples currently under investigation are:

- Grain growth in pure (polycrystal) nickel [4.3,4.4].
- Mapping polycrystal microstructure in a Ni-based superalloy around a fatigue crack.
- Mapping evolution in plastically deformed microstructures in tensile tests of pure copper and zirconium [4.4].
- Measurement of recrystallization and recovery in pure aluminum [4.5].

The grain growth experiment is the first such measurement of its kind and, for example, enables us to measure the effect of anisotropy in boundary energy with predictions of MacPherson–Srolovitz theory. The mapping of microstructure around fatigue cracks permits us to model the deformation associated with crack propagation and be fully informed about the microstructure. Similarly the availability of undeformed and deformed 3D microstructures in plastic deformation allows us both to characterize and to compute them with none of the ambiguities inherent in conventional metallography. A recent overview of the state of the art in this area is found in Ref. [4.6].

4.2.2. Demonstration of chemical imaging virtual experiment

Full-field transmission X-ray microscopy (TXM) offers spatial imaging of the chemical composition of materials. It exploits tunable, high brightness X-rays from a synchrotron, and Fresnel zone plates to obtain high resolution tomographic images. In particular, TXM permits images to be obtained above and below specific absorption

edges for specific elements. Liu et al. describe how monochromatic X-rays in the range 5-14 keV were used with a zoneplate to produce 3D images of elemental distribution in a sample with ~20μm diameter of a Ni-Cu aluminide [4.7]. Just as for the HEDM technique, a virtual TIXM experiment will help users and potential users assess the likely outcome and gain expertise before arriving at the beamline. Given a 3D model of the proposed sample in the form of a 3D image (see e.g. [4.8]), the absorption could be simulated in enough detail to allow a user to run the tomographic reconstruction software and construct difference images, and obtain element-specific maps. Guided simulations of how to set up the experiment and operate the various pieces of software would be extremely useful to both potential users and occasional practitioners. As before, the workflow would involve users, the software, beam scientists to answer questions about the alignment phase, and expert users to answer questions in the data analysis phase.

For new investigations, TXM could be used to study devices such as batteries, capacitors, and fuel cells, which are multi-material structures whose composition varies in both space and time. Many of these structures have combinations of metals, oxides, pores and other materials that need to be mapped with enough detail that size distributions and dihedral angles between interfaces can be identified, for example. High throughput access to TXM characterization would permit energy storage and degradation mechanisms to be investigated in greater detail.

[4.1] J. Oddershede, S. Schmidt, et al. Materials Characterization 62, 651 (2011).

[4.2] S.F. Li and R.M. Suter, J. Appl. Cryst. 46, 512 (2013).

[4.3] C.M. Hefferan, S.F. Li, J. Lind, U. Lienert, A.D. Rollett, P. Wynblatt, and R.M. Suter, Computers, Materials, & Continua 14, 209 (2009).

[4.4] S.F. Li, J. Lind, C.M. Hefferan, R. Pokharel, U. Lienert, A.D. Rollett, and R.M. Suter, J. Appl. Cryst. 45, 1098 (2012).

[4.5] C.M. Hefferan, J. Lind, S.F. Li, U. Lienert, A.D. Rollett, and R.M. Suter, Acta Mater. 60, 4311 (2012).

[4.6] M. Miller, et al., Synchrotron Radiation News 25, 18 (2012).

[4.7] Y. Liu, et al. Anal. Bioan. Chem. 404, 1297 (2012).

[4.8] D.M. Saylor, J. Fridy, B.S. El-Dasher, K.-Y. Jung, and A.D. Rollett, Metallurgical & Materials Transactions, 35A, 1969 (2004).

5. Uncertainty Quantification

Group leader: S.J.L. Billinge

Group members: M. Aivazis, J. Beck, K. Broersma, A. Cunsolo, P. Dawson, P. Duxbury, R. Ghanem, M. Green, M. McKerns, H. Owhadi, S. Vogel

5.1. Introduction

Uncertainty quantification (UQ) is an emerging field in applied mathematics that aims to put on a more rigorous mathematical footing some of our most important scientific activities [5.1]: estimating uncertain outcomes from imperfect information. This has been a long-standing goal in the physical and biological sciences, and advances in the applied mathematics associated with this objective have revolutionized what we know about the world, for example, with the rigorous formulation of the mathematics of statistics in the 1930s and beginning in the 1940s the development of information theory [5.2]. The goal of current UQ research is to provide the requisite rigor and procedures for analytic certification, which is to certify the credibility of analytical knowledge, in support of decision-making [5.3]. It aims to go beyond the notion, as powerful as it is, of estimating the average value of a stochastic variable, and the uncertainty of that average value, to tasks such as placing bounds on predictions of extreme events and determining the reliability of predictions. We see an important role for UQ in the scattering sciences, particularly in the area of solving chronically ill-posed scattering inverse problems such as (but not limited to) the nanostructure problem.

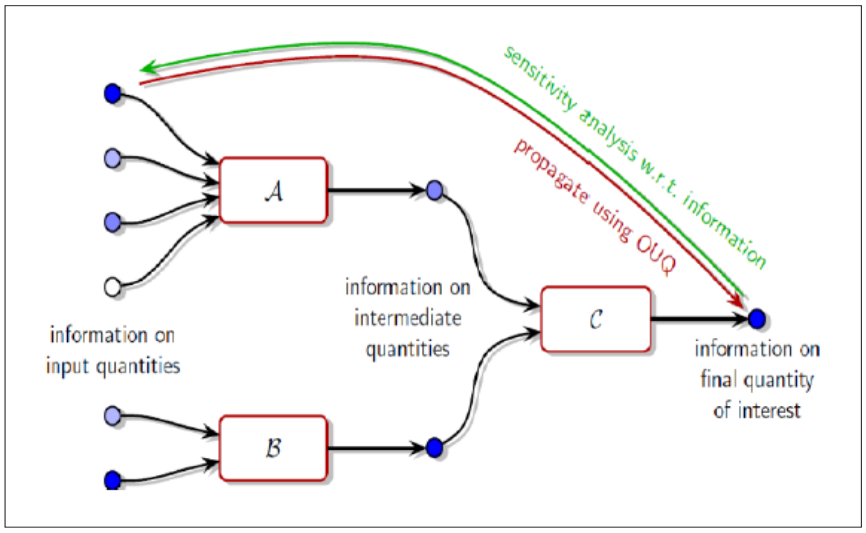


Fig. 5.1. Schematic of how UQ fits into an inverse problem workflow. UQ provides machinery for optimally determining the model from the prior information (red curve). It is also able to place bounds on our uncertainty of the model parameters (sensitivity) given the initial information (green curve)

5.2. Overview of Uncertainty Quantification (UQ)

UQ is the business of computing optimal bounds on quantities of interest given the information at hand. UQ can be logically divided into three groups:

- Inverse problem whereby statistical characterization of input variables are estimated
- Uncertainty propagation whereby the statistical characterization of output variables obtained are constrained by the physics of the problem and the statistics of the input variables
- Error attribution whereby a worth of uncertainty reduction is estimated for all the uncertainties included in the analysis.

Once a UQ analysis is conducted, a number of tasks can be carried out in a post-processing phase that include model validation, model ranking, and model update either to achieve validation or to assimilate new knowledge in the form of data or physical insight [5.4]. In the present era of big data, big computers and complexity, relevant decisions are often related to analytical knowledge through large scale simulations and so there is a significant computational component to the application of UQ to our materials and scattering problems.

At present, UQ is also an umbrella term that encompasses a large spectrum of methods. It is

an emerging field of applied mathematics, lacking a coherent general presentation, much like the state of probability theory before its rigorous formulation by Kolmogorov in the 1930s [5.1]. This presents challenges and opportunities for our scattering software initiative. However, the opportunities far outweigh the challenges. Advances in UQ take place in response to the needs of scientific (and more broadly, societal) communities; by tightly coupling a UQ component to the software institute, we have the opportunity to steer UQ developments towards the specific problems of the scattering community. These developments can then be made available rapidly to the broader materials science community through the facility user programs and programs of the institute. However, UQ currently is sufficiently well developed that it is already having an impact in a number of areas, for example, to certify or de-certify the safety of complex systems [5.1]. One of the powerful outcomes of UQ is the ability to place rigorous bounds on quantities of interest given the information at hand, allowing us to validate and improve models and to design optimal experiments. This kind of approach is particularly important for expensive forward problems, and it is hard to think of more expensive forward problems than actual scattering experiments.

The immediate application in scattering science will be to quantify the robustness of predictions made with respect to assumptions of the model and the lack of information in inverse problems. This allows for identification of critical variables, assumptions and vulnerabilities of the model. This analysis can sometimes give surprising and non-intuitive insights helping to identify the critical variables and critical missing information needed to regularize and improve a model, guiding us towards the missing experiments or information that lie "on the critical path" for solving the model.

Related is the selection of the most decisive and most predictive additional experiments. This aspect will be enormously valuable for optimizing the use of experimental time, not only making the best use of expensive and valuable resources (including the time of skilled personnel) but will allow the scientist to move much more rapidly to a successful scientific outcome, without months or years of delay. It will also allow us to assess the performance and the design of complex systems in a minimal number of experiments (with predicted bounds on the total number of required experiments). Just as a good chess player thinks several moves ahead, UQ allows for the design of increasingly sophisticated and optimal sequences of experiments that can be performed to measure key system variables. The longer-term vision is that UQ combined with high performance computing will enable the scientific computation of optimal or nearly optimal models.

Another important contribution to the inverse problem is information coming from theory. In general, the inverse problem needs input from all prior information in the form of experimental results, but also theoretical results. The theoretical results discussed in the rest of this report are really forward problems and in the context of the inverse problem they will have to be computed many times in a regression loop. The software developments on the theory side of the institute will be heavily leveraged by the use of UQ. Again, these forward problems are computationally very expensive and it will be critical to use UQ to assess which calculations are likely to provide the most information; this approach will constrain the problem so that we can focus on carrying out the smallest number of calculations and still answer our scientific question.

5.3. UQ in the Scattering Institute

The intent is to make UQ part of the culture in scattering studies of materials. To that end we plan to:

- Determine the best way to formulate questions/ quantities of interest (QoI) for scattering use cases that are amenable to rigorous analysis using UQ
- Quantify the robustness of predictions based on the information (or lack thereof)
  a. Establish robust quantitative ranking of different models given the data and state of knowledge and prior information.
  b. Identify the best place to improve a model. The computation of optimal bounds on the accuracy of a model allows for the identification of vulnerabilities and major sources of uncertainty (via the identification of active constraints/information at the extrema).
  c. Identify experiments that will provide the

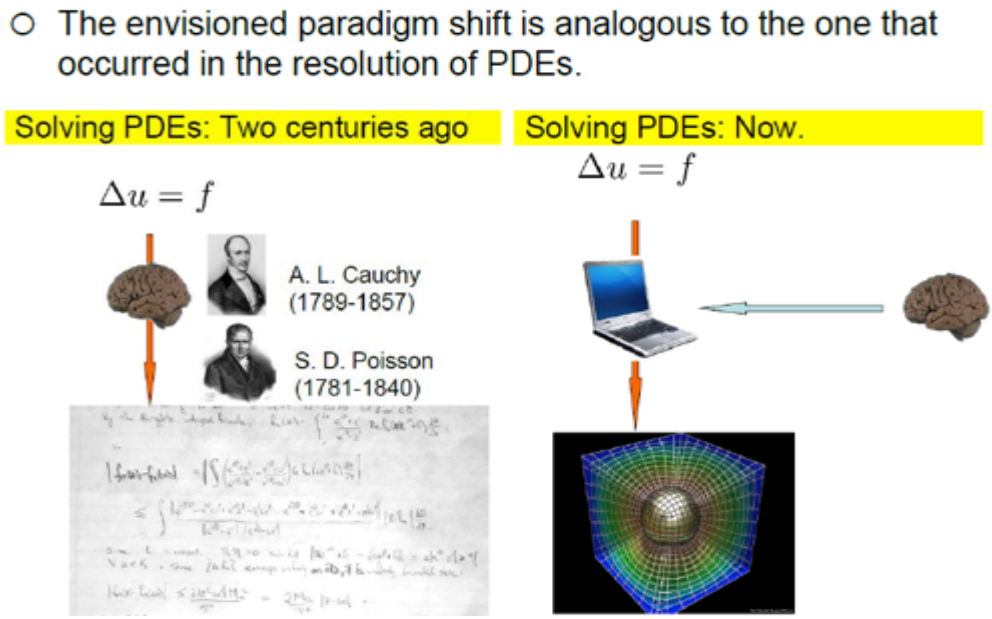


Fig. 5.2. Illustrating the paradigm shift in solving PDE's made possible with advanced computing.

most complimentary information towards answering the QoI. Optimal experiments can be identified as those targeting vulnerabilities (as identified above) or those leading to the maximal reduction in the prediction interval of confidence in the worst or most likely case with respect to the possible outcomes of those experiments.

d. Surrogate modeling => validating approximate but fast models that capture the right physics of slow, computationally intensive, calculations. By ranking models in terms of accuracy and computational cost one can design optimal model selection strategies given limited computational resources.

e. Determine "error bars" or uncertainties on theoretical predictions. Partition them onto different parts of the theoretical model to understand the different sources of error in the predictions.

Longer-term visions include

- Information gaming. Allow a user to test different information sources in a virtual environment to see their effects on the outcome. One long-term vision is to achieve the capability of modifying not only model parameters but also critical model assumptions on the fly to see the effects of those assumptions and pieces of information on quantities of interest and their uncertainties.
- "Current Optimum Model" (CoM). As new information (e.g., experimental results) become available they are fed into the CoM program, which updates to yield a new CoM dynamically. The idea is that the optimal model is evolving as new information comes in (and is validated). This optimal model could be open source and evolving with the input of the community as a whole (an analogy would be the evolution of Linux/Unix).

5.4. Workflows

We see that rigorous UQ principles, and software implementations and expertise in the institute, will have significant effects on workflows of scattering scientists, from experiment design, to optimization

of resources (including beamtime), to handling raw data, to modeling and model validation, and feeding back to the design of new complementary experiments. As the tools develop we expect that the sophistication of the questions that scientists can ask, and design suites of experiments to answer, will become ever more complex and powerful.

The institute would actively explore the ways in which UQ can add value to scientific problems in the domain of scattering and materials science. We will identify what is missing from current UQ frameworks and understanding, and therefore what UQ research needs to be done to make it useful for our particular scattering problems. As with software development in general, this is best done by prototyping. In this context, this means finding a particular, representative, science problem and attempting a solution with UQ, while keeping in mind that we want a solution that is flexible enough to be repurposed to other related problems in the future. We have identified the scattering inverse problem in the form of solving nanostructure from powder pair distribution function analysis data as our best candidate for a prototype system. When this is successfully demonstrated, we will develop a robust and flexible software framework for applying the UQ approach more broadly. The workflow will then involve scattering scientists and materials scientists interacting with the UQ framework tools during experiment conception, data gathering, data analysis and interpretation stages. The software will help to predict the best candidate experiments to carry out, and fragilities in the analysis that need additional information. It is therefore an iterative process of asking a question, collecting data, analyzing and modeling the data, assessing the validity of the answer to the question, determining the best next experiment, carrying out that experiment and adding the experimental information to the inverse problem, reanalyzing and reassessing the results, and so on, until we have answered the question to our satisfaction.

[5.1] Optimal Uncertainty Quantification. Houman Owhadi, Clint Scovel, Tim Sullivan, Michael McKerns and Michael Ortiz. SIAM Review Vol. 55, No. 2 : pp. 271-345, 2013 (Expository Research Papers)

[5.2] R. E. Barlow and F. Proschan. Mathematical Theory of Reliability. SIAM, Philadelphia, 1996.

[5.3] I. Elishakoff and M. Ohsaki. Optimization and Anti-Optimization of Structures Under Uncertainty. World Scientific, London, 2010.

[5.4] J. L. Beck and L. S. Katafygiotis. Updating models and their uncertainties: Bayesian statistical framework. Journal of Engineering Mechanics, 124(4):455–461, 1998.

6. Software and Hardware

6.1. Software development

Group leader: T. Proffen

Group members: E. Deelman, B. Fultz, M. Hagen, J.Y.Y. Lin, L. Lutterotti, M. Newville, R. Neder, J.J. Rehr, D. Roach, G. Shipman, M. Stalzer, M. Tucker, L. Udby, M. Wilde

6.1.1. Skills and methods for software projects

Software projects will be regular activities of any software institute, and these projects need to be managed to maintain schedule, to utilize appropriate resources, and, first and foremost, to create useful and robust products. Cost, schedule and quality are risks for any project, and are particularly challenging for software projects. Historically, a high risk has been insufficient communication of the user needs to the software developers. Intense attention to requirements is necessary when the software developers are not experts in the domain of the user, and the users are not experts in software development. The situation is more fortunate for scientific software development when the developers are also scientists themselves who understand computational scattering science. Such persons are essential to a software institute for scattering science, and the skills of these individuals and their guidance will determine the success of new software developed by an institute. Because so much depends on the uniqueness of individuals, it is important to consider their skills and motivations. Career paths are discussed in section 7.

There are relatively few persons today with professional skills both in the hard sciences and in software development. Many have found the opportunities for computing in the sciences, and many others have seen how software design contributes to high performance computing. Relatively few have done both, and fewer still have done both in the field of computational scattering science. Furthermore, scattering science encompasses a wide range of kinds of scientific disciplines, e.g., condensed matter physics, chemistry, crystallography and thermodynamics. This brings unique challenges to software development for scattering science. Unlike scientific disciplines with one core computing problem (such as solving a class of PDEs like the diffusion equation or the wave equation), the software packages for scattering science contain diverse types of underlying mathematics including, just to name a few, FFT, symmetry analysis, optimization, Bayesian analysis, and Monte Carlo methods. Many of these mathematical problems have software implementations in other scientific domains, and their main task could be regarded as being as simple as wiring different algorithms together. The real challenge, however, is expressing the scientific models in mathematics for algorithmic implementation, and in translating to computational representations. This needs to be done in a robust way that allows future adaptation to new problems in computational science. In addition, there are problems where research on new tools is necessary.

In spite of this diversity of needs and tools for computational scattering science, the sharing of software infrastructure such as source repository, automatic building and testing, and code review tools would prove useful to all teams of a software institute and the whole scattering science community outside the institute. A common set of development practices can be promoted with such a community resource, and some suggestions are given in what follows.

6.1.2. "Lean startup" principles

A central focus on scientific outcomes is required for the success of a scientific software project. Without such a core mission, and because software development can be intellectually challenging and fun for its own sake, projects can go out of control and drift from the main scientific goal. Strong leadership and good management are needed. Although professional software engineering in the commercial sector has been ongoing for decades, and there is extensive experience in the software industry from which we can learn, the development of scientific software is unique. Experiences from scientific software projects are more relevant for project planning in a future institute.

We suggest that software development for computational scattering science should be like that of a startup company. As in startups, scientific software developers are undertaking highly creative activities with very limited resources. Resources

must be invested wisely, and not wasted on rigid management tasks or fruitless abstract discussions. In particular, under rigid procedures of planning, designing, testing and implementation cycles, it may take years for the team to develop a first version of a software package. The package will be hardened and robust, but what if users are not happy with this end product? Since resources are scarce, an institute needs a more agile approach to software development projects, such as "lean startup," defined in Wikipedia as: "an approach for launching businesses and products that relies on validated learning, scientific experimentation, and iterative product releases to shorten product development cycles, measure progress, and gain valuable customer feedback."

Not all practices of lean startup are relevant because it is geared towards commercial applications, but some principles can be adapted to scientific software engineering. In particular, the core idea is simple yet powerful: working backwards from what scientists want, instead of working forward from what technologies can do. A development process with a lean startup philosophy will first start with an idea of a software workflow (or perhaps a feature in it), decide on a small set of deliverables, and then implement the first prototype as quickly as possible. The prototype may be even as simple as, for example, a picture of a user interface. The goal is to get user feedback early. Modifications to the idea are made accordingly – sometimes the first idea needs to be thrown away, but little has been invested at this point. With a better idea of what users want, developers build another prototype, using tools that are most efficient. Only after users confirm the usefulness of the software (features) to scientific research should more resources be spent to rethink and improve the design, optimize performance, and build better user interfaces. In this kind of development process, continuous communication between developers and users and fast iterations are key. This engagement takes real work by users, and will not happen without users having personal commitment to the effort.

A software institute can ensure this investment by engaging the user scientists as collaborators in a project that will soon lead to publishable scientific results. After this collaboration, the workflow design and quality need assessment, and either this assessment or a revised software workflow needs to be archived.

6.1.3. Sustainable software

Emphasizing a quick development cycle does not necessarily mean sacrificing quality for schedule. Sustainable software can still be built by following a few practices that have been found especially suitable for scientific software engineering.

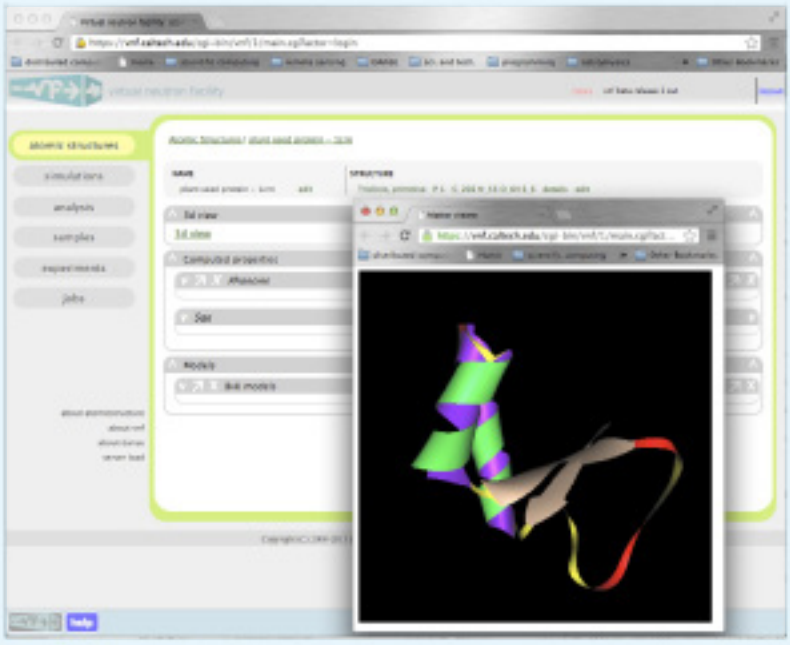
• Simpler is better. Simple design and implementation are easier for developers to maintain, debug, and improve.

• Use high-level languages whenever possible. High-level languages such as Python lead to cleaner, more readable, hence more maintainable code. Code performance should be optimized by using "smarter" constructs, or by using implementations of lower languages only if a piece of code becomes a performance bottleneck. This way, when new high performance computing techniques are developed, the codes are easier to understand, are more flexible, and can evolve to catch up with new technology.

• Implement scientific models in a structural and general way whenever possible. Scientific algorithms are already hard enough to understand. If the code that implements a model is written in a mysterious and non-transparent way, it is hard to imagine how it can be maintained. It is better if the mathematical and programming representation reflects the structure of the scientific model in a general way, with appropriate documentation that explains the big picture and its substructure.

• Use design patterns judiciously. Object-oriented design patterns are validated, reusable solutions to programming problems. When used correctly, design patterns can help developers create solid code and better documentation. Design patterns can be overkill, however. If misused, design patterns can produce overly-complicated code that takes longer to develop, is harder to explain, and is more difficult to maintain.

• Use component architecture to compose programs from known working elements. These components can often be built with an eye towards reusing them in future computational workflows.



Screenshot of the Virtual Neutron Facility. The VNF is a web application that integrates scientific software packages for material simulations and Monte-Carlo simulation of neutron scattering. (J. Y. Y. Lin)

• Follow coding standards, use inline documentation, and develop component level regression tests. These words cannot be given as platitudes – an institute needs to embed them in its culture.

The guidelines above are for individual developers. To have a cohesive team, communications among developers are essential, and these interactions set the style and work in an institute. Two important technical practices are:

• Code reviews. These should occur at critical stages, such as for a baseline design and before release. The reviews are easier if a developer has planned documentation and has cleaned up shortcuts as a regular work practice, and encouraging these practices is a benefit of code reviews. The other developers on the review should expect to invest some time themselves, as they would expect for reviews of their own code. It is usually appropriate to confine the reviews to developers themselves. This allows better focus on technical issues, and makes the review less intimidating to new developers who can learn quickly by participating in code reviews. Informal code reviews are also valuable for product development and professional growth.

• Code sprints. The software institute will probably have a geographical distribution of developers. It is useful for these teams to gather together a few times a year to do code sprints. Writing code together can help developers deal with subtle and complex coding problems, and "synchronize" developments across the institute. Developers in different teams can exchange and enhance their ideas, skills, and codes, and plan for changes in design.

These practices require balance between ideology and practicality. It is dangerous to be a perfectionist in one's own work, or intolerant of imperfections in the work of others.

6.1.4. Development team

Scientific software development requires two types of mindsets from its developers: engineer and scientist. Good researchers who lack craftsmanship will write

poor code. Good coders lacking genuine curiosity of science will work for commercial firms that pay better. It is challenging but absolutely essential to find the right people for the development team. With risk it is possible for computational scientists to take on software engineering tasks, and sometimes vice-versa, but there is a time needed for this transition. Considering either the transition time or the effort to conduct a job search or both, any major changes in staffing take nearly a year to equilibrate. A permanent or semi-permanent staff is necessary for steady operations.

The previous subsection described roles and responsibilities for the scientist/software developers who are essential to the success of an institute. Section 7 gives more information on a staffing plan, but roles for other key personnel are mentioned here. An essential member of a development team is the software architect, who is responsible for designing the software systems. A good architect also teaches good ways of thinking about programming, and helps train developers in best practices for building sustainable software. Code reviews by the architect are most helpful, since these teach best practices by example, and are important for building more sustainable software systems. For the development of scientific software, domain experts should be included in the team and supported as a part of the team. They can be graduate students and postdocs who help the software development by using the software and giving continuous feedback. For service to the community, applications specialists are needed for supporting users. They should have some software development skills, but mainly they are helping users doing their research, similar to the role of instrument scientists in national scattering facilities. Coordinating the community interactions is a job for a "catalyst," who would help the institute seek out opportunities for scientific collaborations.

System administrators are needed for setting up and maintaining the hardware and system software for servers and high performance computers. A developer is often needed as an infrastructure builder. He or she builds the supporting infrastructure such as the code repository, automatic building and testing system, and code review system.

Repeating a key point, the software for scattering science is diverse. There are, however, two main categories of software:

- Software that is more focused on user interactivity
- Software that needs high performance computing resources.

Development work in these two categories requires different skill sets. For software more focused on interactivity, the developers need to work quickly and continuously with users to achieve the best usability. For software that needs massive computing resources, the emphasis on is on the developers' ability to optimize and parallelize the code to achieve better efficiency. Hence, at least one specialist on quick prototyping and GUI development and one specialist on performance optimization and parallelization (thread parallelization, MPI parallelization, GPU computing, and hybrid parallelization) are needed in the team.

Returning to the computational scientists on the team, an institute must be able to keep them happy. To do so, it is essential that they take pride in the scientific output generated from the software they have produced, since salary in academia is not comparable to that in industry. The way to do this is to engage them in the creative process of doing science, credit them with published scientific results, and involve them as best as possible in the mission of the university or national laboratory.

6.2. Software Technical Approach

Owing to the diversity of scientific needs in computational scattering science, it is unreasonable to develop all codes in one large application, or even under one tightly integrated framework. The focus instead should be on ensuring that the data objects and high-level APIs of different computational components can be coupled into useful workflows, and that these workflows can be maintained and developed further in the future. This implies some balance between present and future workloads, and such decisions should be the responsibility of an institute staff and its oversight board. Although detailed technical decisions must be left to the institute and its staff, some general practices can be stated now.

6.2.1. Glue language: Python

At the lowest level, scientific codes were, are, and will be, written in several different languages, with

the most popular being Fortran, C, and C++. These codes, which are validated and understood, cannot be rewritten substantially without loss of community trust. (There are, however, codes written for proprietary computing environments such as Matlab, IDL or Mathematica, which tend to be shorter, and perhaps should be rewritten.) A high-level language can wrap existing and future codes, and bind them together in workflows. Python has proven its success in that regard both in scientific computing in general and in scattering sciences specifically. Other candidates include tcl (not object oriented), perl (hard to maintain when a project grows), and julia (very promising, need to watch its development).

6.2.2. Workflow

Combining components into workflows becomes relatively easy when all components are well defined and implemented. In the early phase of building a workflow and gathering user requirements, simply using Python to glue together modules is an efficient first step. After a workflow is built and tested, and when it is time to prepare to launch workflows over a large number of resources, SWIFT may be considered for workflow construction. If users want to visualize and customize their workflows, VisTrails may be useful for implementation.

6.2.3. Build software components with clear API, documentation, and tests

Software components must be designed with interoperability in mind. We believe this is the single most important criterion for building a sustainable software infrastructure. Clear and robust component interfaces and solid implementations are key to building workflows that last, and can be modified and maintained into the future. It would be most helpful if all codes within the institute were to use the same component-based framework that helps teach and enforce the related principles of modern software design. It is challenging to enforce strict rules on scientific software developers. The architect and the lead developers must expect some rebelliousness, and have the wisdom to know when to tolerate it.

6.2.4. Achieve performance with care

An important rule of thumb is "simpler is (usually) better." A common pitfall for a scientific programmer is inventing "smart" and sophisticated data structures or algorithms that prove to be no better than a much simpler implementation in performance or resource use, but are harder to maintain. In scientific computing it is usually wise to trade speed for clarity. For example we should

- prefer high-level languages like Python to low-level languages like C,
- prefer OpenMP to direct thread manipulation,
- defer performance optimization until after
  -- the usefulness of the program is proved, and
  -- scientific research is known to benefit from the optimization task.

Sometimes it is more effective to trade CPU hours for the time needed to optimize a code. Furthermore, the lifetime of the software products from the institute is expected to be longer than the lifetime of the computing infrastructure. It is therefore more valuable to focus on the higher-level computational representation of the physics and math problems, instead of the details of low-level optimization.

An institute should, however, use and contribute to open standards for high performance computing such as OpenMP, MPI, OpenACC, to allow the institute to rely on them to produce maintainable, high-performance code. Meanwhile, Chapel is a promising parallel computing language that provides a generic parallel programming model and concise syntax. We should watch its development.

6.2.5. Data I/O

Data size and rate in scattering science, although increasing quickly, are far smaller than what is common in high-energy physics. The main computational tasks on data I/O for the Institute are to define data models and data formats. Sometimes this may require building a consensus within the community, or at least engaging the community in the decision. In other cases, libraries such as HDF and netCDF should be used without discussion.

6.2.6. Graphical user interface

Graphical user interfaces are becoming an expectation for users of scientific software. Python-based native user interface construction is time-consuming but relatively straightforward. Care should be taken to promote a common look and feel among the native GUIs created by the Institute, and this can be achieved by establishing guidelines for native GUI construction. Web-based interfaces for scientific applications are gaining momentum, but they are often designed poorly. LubanUI, which provides a high-level abstraction for programming web-based user interfaces, should be considered because it hides the complexities of web programming languages and widgets, and simplifies maintenance and evolution.

6.2.7. Software as a service

At the highest level, software workflows should be available as user services. This makes the software easier to access, easier for developers to upgrade, easier to deploy, and easier for workflow builders to integrate tools into workflows. The Caltech group is developing an approach to exposing functionalities as services that are easily coupled to online data visualization tools. Globusonline may be appropriate for file transformation and user management for software services. HubZero seems appropriate for hosting web presentations of the software packages and workflows.

6.3. Use Case

Several types of software projects are likely to be regular activities in a software institute, including

- Integrating software components into a new computational workflow, or modifying an existing workflow to do new science.
- Developing completely new code to accommodate developments in theory or experiment.
- Modernizing an existing code, for example to allow it to run well on new hardware architecture.
- Building a graphical user interfaces for a command line code.
- Developing libraries and common tools aiding development.

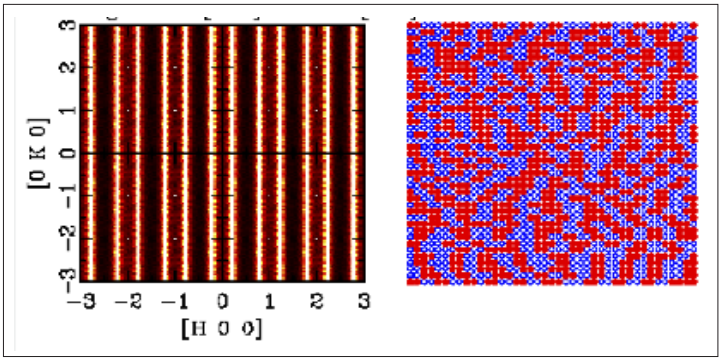
As an example, this section outlines the use case of modernizing the DISCUS code (http://discus.sourceforge.net). DISCUS is a program to simulate disordered materials and calculate related scattering quantities such as the powder diffraction pattern or single crystal diffuse scattering. The program has been developed over the last 20 years and will benefit greatly from being 'modernized'. In the following sections, we outline some essential steps in a DISCUS modernization project. It should be noted, however, that the development procedure should not be regarded as a linear sequence of steps. Instead, it should involve quick iteration of some of these steps, following the Lean Startup philosophy. The example that follows would be a substantial project, and would require administrative approval and oversight.

6.3.1. Brainstorming phase

Extended brainstorming between domain scientists and computational scientists is essential to the success of any software project. In fact it might be the most critical step as it usually takes time for the team to even speak the "same language." A facilitator – a person with sufficient knowledge of the underlying science and of software design can help bridge the gap between the various experts on the project. The initial brainstorming will be focused on the requirements and scope and result in the outline of a prototype to be developed.

6.3.2. Analyze existing code and establish requirements

The first step will be a detailed analysis of the existing DISCUS software package, which includes



DISCUS - diffuse scattering and structure simulation - showing single crystal diffraction pattern. (R. Neder)

the DISCUS code itself, along with helper applications for refinement (DIFFEV) and plotting (KUPLOT). The initial product of this stage is a list of features, algorithms, supported file formats and so on. This is the foundation for the requirements for the new code. A decision needs to be made about which parts of the original code will be kept, rewritten or replaced with other existing utilities or scientific libraries not currently in use. Other requirements include input/output data structures and file formats including the questions of legacy format support. This is also the time to determine if the code will benefit from the ability to scale to HPC platforms. Meetings with current and potential users of the DISCUS package should also help determine requirements, which will become the road map for the project.

6.3.3. Define scope

Although the requirement gathering might answer this question in part, it is essential to establish and agree on the scope of the modernized code, which can be the same as the original code or extended as new features are added. The scope needs to define the boundaries as well as desired interactions with other existing or future code. The scope should be subject to change, however, as iterations of development—user-validation cycle will help refine and even change the scope of the software package. In the case of DISCUS, for example, the scope might exclude visualization and simply require output in standard formats to interface with external visualization packages, e.g. structure viewers.

6.3.4. Prototype development

After requirements and scope have been established, a first prototype of the modernized tool is developed. The prototype might implement only a subset of the features. Possible migration to HPC platforms might not be included at this step as this is intended as an iterative step. It is important, however, that the initial prototype includes enough features to allow testing and validation by the domain scientist and external "friendly" users. Feedback from testing, fixing of bugs and growing the functionality of subsequent prototypes will be the main goals. In fact the tasks of 6.2.1, 6.2.2, 6.2.3, and 6.2.4 will certainly form a cycle leading to a more and more complete prototype.

6.3.5. Production

After a fully functional and tested prototype of the modernized code has been developed, it will need to transition to a production version. An automatic testing infrastructure for regular testing of the software on all target platforms (e.g. Windows, Linux and/or Mac), should be available as part of the institute infrastructure. As the release date approaches, binary versions and automatic installers need to be created and tested on all target platforms. In addition to developer documentation of the code and algorithms, user documentation and tutorial examples must be available and reviewed by a potential user. An important step, often overlooked, is to create a website, workshop report or scientific paper that can be cited by scientists using the code. At this stage the project can be considered finished, and a modernized version of DISCUS declared available.

After release, a critical need will be code maintenance including bug fixes and responding to user questions and requests. In many cases (e.g. large open source projects), there is sufficient knowledge in the community that the latter part can be largely handled by forums, blogs or mailing lists. Scientific software packages usually have a much smaller user base, and an even smaller pool of capable developers. Common facilities that make it easier for "occasional developers" to contribute to the software projects prove useful if quality can be assured. The software institute should explore platforms and policies for user contributions.

Once again, we emphasize the importance of fast iterations in development. To shorten the turnaround time of development iterations, the production and release phases should begin early. For example, prototype releases should be handed out to early adopters of the new software for feedback early and often, and an automatic building and testing system (at least a primitive one) should be available from the beginning of code development. A website for the software project should be established as early as possible, to facilitate easier communication not only among developers, but also between developers and early users.

6.4. Hardware Trends in Computing

Group leader: M. Stalzer

Group members: S. Brunett, M. Graham, P. Messina, J. Muñoz, D. Roach, G. Shipman

A five-year lifetime of the Institute will see computing systems evolve for both simulation and data-intensive workloads. This section describes our understanding of these trends and the opportunities for SIXNS.

6.4.1. Systems for simulation

For simulations, the most significant trend in computing is the dramatic increase in parallelism on a single processor socket. Ten years ago processors executed a single instruction stream, but in 2013 it is not uncommon to have socket-level processors that execute 256 streams simultaneously. The Intel Xeon Phi shown in Fig. 6.1 is an example. Each core of the Xeon Phi can execute four hardware threads simultaneously, typically two down both its vector and scalar pipes. These threads have access to a L1 cache of 32KB and an L2 cache of 512KB. The L2 caches across all 64 cores in the on-socket ring are coherent, meaning that writes to the L2 cache of one core

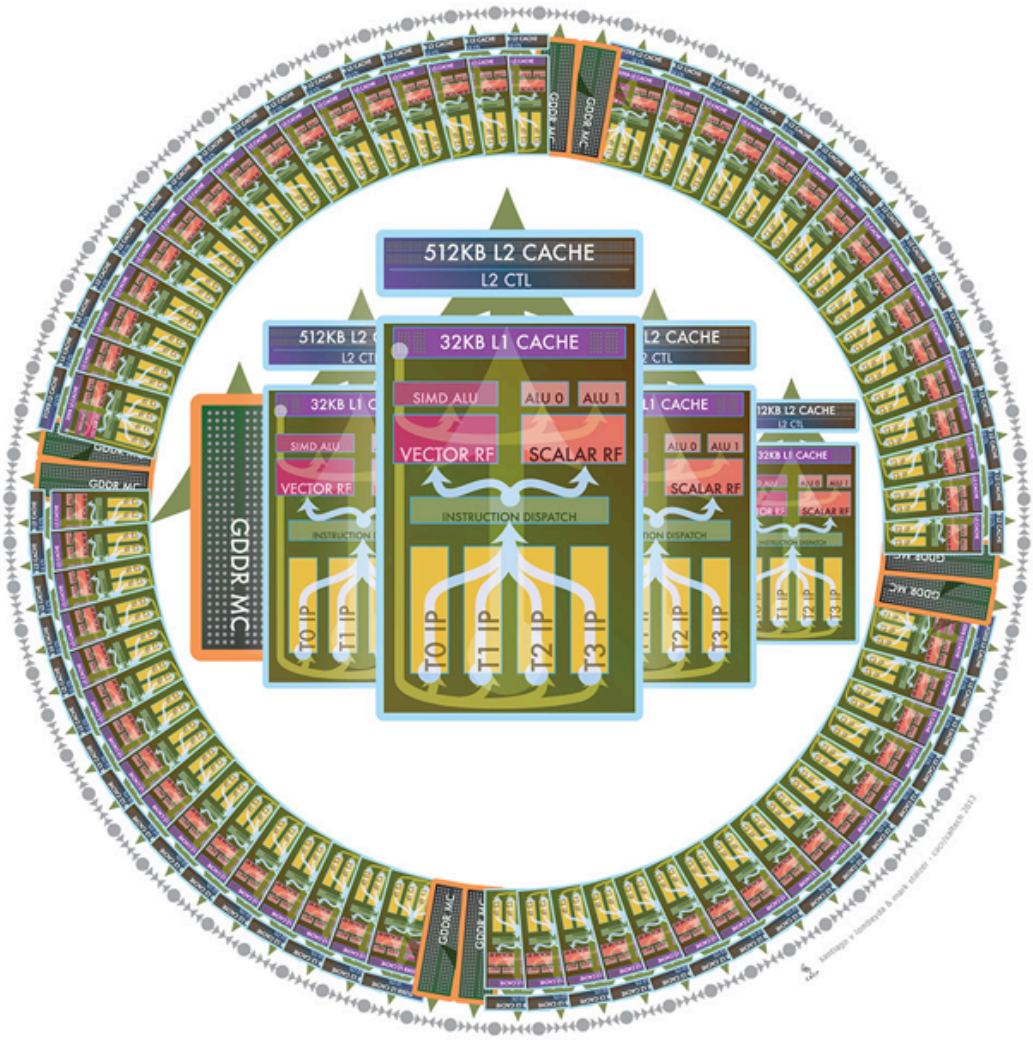


*Fig. 6.1. Architecture of the Intel Xeon Phi. A single socket contains up to 64 cores and is capable of a teraflops. Note that there are 4*64=256 parallel threads on a socket. Besides the 64 cores, there are 8 controllers to off-socket memories (GDDR MC) and a PCIe link (not shown) to the host processor.*

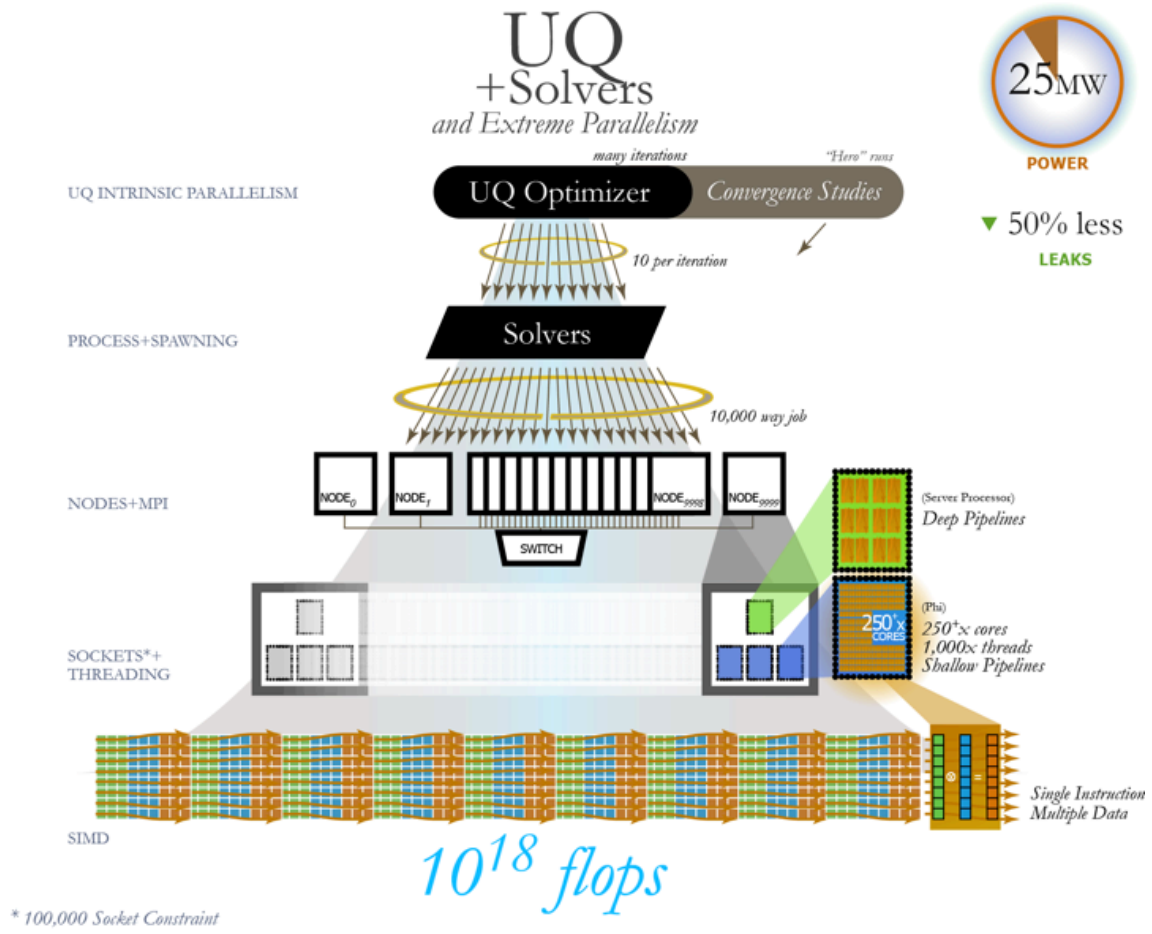


Fig. 6.2. The layers of parallelism of an uncertainty quantification campaign at exascale (c. 2020). The processor in Figure 6.1, evolved over ~5-7 years, is one of about 100,000 of the elements at the bottom of the pyramid.

are seen by the other cores' caches. The chip looks like a complete symmetric multiprocessor with distributed shared memory. There is a further level of parallelism within a core – one of the two execution pipes is a vector unit capable of completing eight double precision multiply-accumulates per clock cycle. At one GHz, the performance is 109*64*8*2 ≈ 1012, or a teraflop. The Xeon Phi sustains this performance on the Linpack benchmark. Programming such a device for peak performance is not easy, and particular care must be taken in staging data into the caches to maximize temporal locality.

In general, it is not enough to "tune" the simulators to take advantage of such an advanced socket level architecture: the inner loops and data layouts must be redone.

The largest computers in ~5-7 years will comprise approximately 100,000 Xeon Phi-like processors (Nvidia will likely be making a competitive product), giving 1018 flops (an exaflop) under optimal circumstances. This will add more layers to the parallelism, in a sort of pyramid as shown in Fig. 6.2. At the socket level we have threads and vectors (SIMD). Then we move up to process parallelism using MPI of rank ~10,000. At the highest level is a UQ process that typically launches about 10 jobs at once to explore parameter space, feeding the entire machine. Another mode of operation is a "hero" run used for convergence studies. Here a single 100,000 rank job occupies the machine and there will be several hundred million threads working on the same calculation. Another point about Fig. 6.2 is the projected electrical power of exascale machines – 25 MW at minimum, and this assumes a factor-of-two decrease in power consumption per flop.

Some simulation codes are already taking advantage of significant parallelism. GULP's new neutron scattering functionality demonstrated linear scaling on over 2000 cores in early tests

[6.1] and is capable of taking advantage of GPU accelerators to solve millions of Hermitian Eigensystems in heterogeneous architectures. The latter is achieved by leveraging both custom-written code and MAGMA: an ongoing port of the BLAS and LAPACK numerical libraries to multicore processors and GPU accelerators [6.2]. Another example is LAMMPS, which runs with near perfect scalability out to 100,000+ cores [6.3]. While good progress has been made on the effective parallelization of some simulation, other important codes have seen poor scaling on parallel architectures.

These hardware trends present many opportunities for an Institute to serve the scattering community:

- SIXNS can provide consulting services to simulator developers to help increase performance.
- SIXNS can provide an over-the-horizon look at emerging computer architectures, particularly at the socket level, to help developers design and develop new simulators, perhaps with significantly new modeling capabilities.
- SIXNS can provide moderate sized, but advanced, machines for developers to try out and tune codes. These machines would act as staging resources for much larger national resources, thereby increasing the effective utilization of these resources.
- SIXNS can act as an interface between the scattering community and device designers such as Intel and Nvidia.

6.4.2. Systems for data-intensive science

The rapid advancement of high-performance computing for simulation has been due to a combination of Moore's Law and an easy to understand benchmark – Linpack, which is the basis for rating machines in the TOP500 list. There exists no similar benchmark for data intensive science, and these architectures are at least five years behind. The large-scale systems of today are highly unbalanced when I/O and memory ratios are considered. One ratio is the Amdahl #: the number of bits of sequential I/O per second divided by the number of instructions executed per second. There are similar ratios for memory per core (memory) and random I/O operations per second (IOPS). It is not unusual for simulation codes to have an Amdahl # of 10^{-5}; but data intensive applications and scientific workflows might need an Amdahl # of about 1, and this requires significant architectural changes in the computer system.

Similarly, random access I/O (IOPS) is important for data intensive systems. Latency to disk arrays is a bottleneck. There have been several efforts using flash memories between RAM and disk to hide some of the latency. The Gordon supercomputer at the San Diego Supercomputer Center closely couples Solid-State Disks (SSDs) based on flash memories with compute nodes [6.4]. A Gordon supernode consists of 64 Intel Sandy Bridge processors, 2 TB of RAM, and 8 TB of SSDs, all connected in a virtual symmetric multiprocessing configuration so that any processor can see the entire RAM of the supernode. Gordon connects 32 of these supernodes into one system. The IOPS ratio of the system to the SSDs is about 10x greater than for a spinning disk. Overall, the system processes data intensive workloads about 5x faster than with just regular disks. It is available via the Extreme Science and Engineering Discovery Environment [6.5]. Another example of good IOPS is the Cyberbricks project [6.6]. A single cyberbrick has a low-power dual core Atom processor, a 16 core GPU, 4 GB or RAM, an SSD disk, and regular spinning disk. Its Amdahl number (to SSD) is 1, and clusters of cyberbricks have been used for data mining applications where data is streamed off of the SSDs into the GPUs. There are more extreme examples, such as FlashBlades, that combine many low power parts to achieve over 100x I/O performance compared to disk-based systems [6.7].

The volume, velocity and variety (VVV) of data in scattering science will all grow over the next five years.

- Volume: The volume of data varies substantially depending on the experiment. With new detectors at the APS, cumulative data rates could be 1 PB/day, and the trend for X-ray measurements is upwards (cf., Fig. 1.1). For SNS it is about 2 TB/day, but for some scattering experiments it is only a few KB/day.

- Velocity: User facilities have a well-defined notion of data velocity, i.e., the cumulative count rate. Instruments are being added and upgraded regularly, and the data velocity is increasing steadily.

- Variety: The variety of data is also increasing, especially as data are aggregated from multiple experiments. A common data structure is not practical.

Longer-range planning is needed to match the expected VVV requirements of scattering science to the systems for data reduction and scientific workflows. The services of an Institute would overlap with those from simulation, namely consulting services, over-the-horizon looks at emerging architectures for data-intensive science, and the provision of staging resources. In addition, SIXNS could establish benchmark workflows that would highlight system balance with the goal of informing HPC acquisition decisions and increasing scientific throughput. SIXNS would not only help the scattering science community, but also have broader impact for other scientific domains and industry, and help inform national policy discussions.

[6.1] M.T. Garba, H. Gonzalez-Velez and D. L. Roach, Intl. Conf. on Higher Performance Computing and Communications (HPCC-10), 509-514, September 2010, Melbourne, Australia.

[6.2] E. Agullo et al., J. Phys. Conf. Ser., 180, 012037 (2009).

[6.3] S.J. Plimpton and A.P. Thompson, Computational aspects of many-body potentials, MRS Bulletin, 37, 513-521, 2012.

[6.4] M.L. Norman and A. Snavely, Proc. 2010 Teragrid Conf., Pittsburgh, PA.

[6.5] www.xsede.org

[6.6] A. Szalay et al., Workshop on Power Aware Computing and Systems (Hotpower '09), October 2009, Big Sky, MT.

[6.7] M. Stalzer, FlashBlades: System architecture and performance analysis. Proc. Second Workshop on Architectures and Systems for Big Data (ASBD '12), June 2012, Portland, OR.

7. Community-Based Software Institute

Group leaders: B. Fultz, S. Vogl

Group members: P. Dera, E. Fahri, M. Graham, M. Green, M. Hagen, M. Hybertsen, K. Jorissen, P. Messina, M. Newville, R. Osborn, G. Shipman, M. Wilde, N. Wilkins-Diehr, A. Wills, L. Young

7.1. Goals

Software Institute for X-ray and Neutron Scattering (SIXNS) should

- Develop expertise in a field of science that underlies X-ray and neutron scattering experiments.
- Engage the scattering community in the opportunities for new science enabled by computing.
- Find relevance to the mission of national user facilities and broader community outside the particular science of the Institute.
- Develop new software workflows, using modern methods to manage components and their interconnections for long-term reusability.

7.1.1. Focus on science

Many experimental techniques are unique to X-rays, or unique to neutrons. Simulations of instruments therefore require experimental details and expertise. On the other hand, the same materials and phenomena are often studied by multiple techniques. Atomic structure is studied by both X-ray and neutron methods. Dynamical excitations in matter are also studied by both methods. Although the energy range for inelastic X-ray scattering tends to be higher than for inelastic neutron scattering, there are overlaps of the two methods around energies for the excitations of phonons and magnons. For early efforts of the Institute, software workflows that benefit both X-ray and neutron scattering may be good choices. Strict overlap would not be a requirement for the selection of a project for Institute backing; however – studies of hydrogen dynamics by neutron scattering or studies of electronic excitations by X-ray scattering are excellent topics, too.

There is, however, a further advantage of computational workflows that support both X-ray and neutron scattering. When both X-ray and neutron data are obtained on the same physical phenomena, discrepancies in the methods and in the underlying models can be used for developing methods of uncertainty quantification (UQ). Early examples could be those for which X-ray, neutron, and simulation data are available for two or more models.

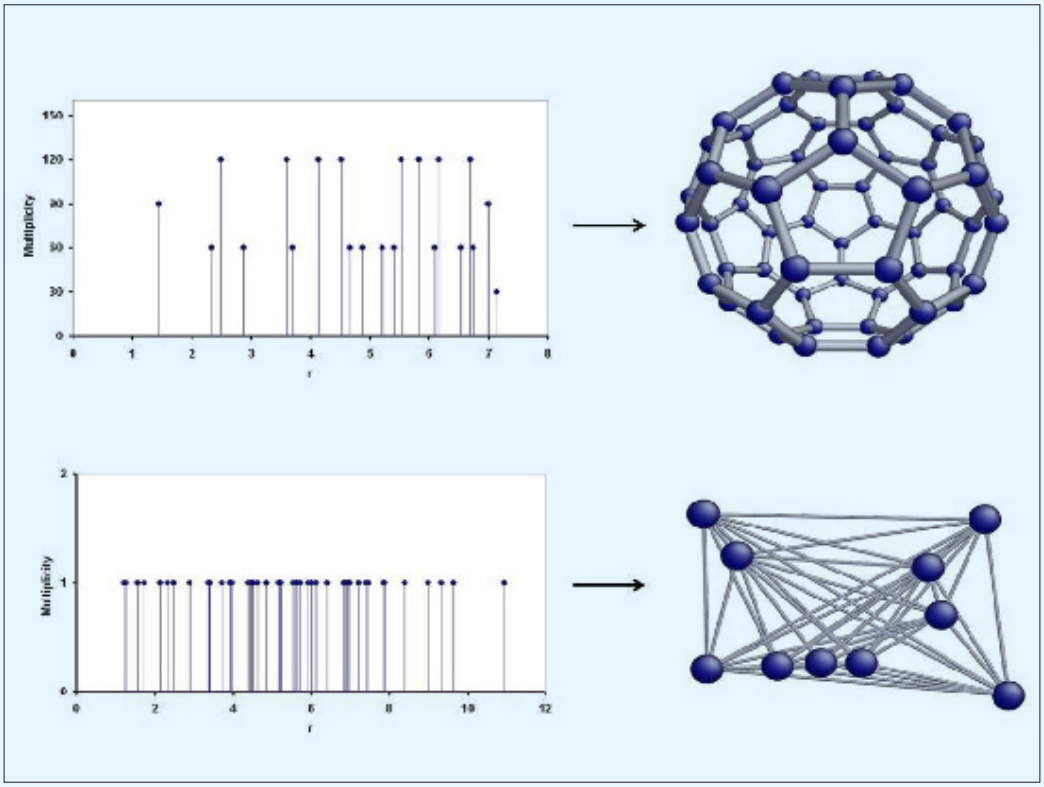
7.1.2. Engage the scattering community

Sections 2-5 of this report presented computational workflows that would elevate scattering science. General opportunities were described, as were some specific use cases that could be supported in the very near future. The Institute would provide the resources and coordination to make them happen, initiating software projects to build the needed workflows. The software projects to build computational workflows would be planned around science, through collaborations between computational scientists and experimental groups. The goals should be typical of scientific collaborations, leveraged by Institute resources. These extra resources would give more opportunity for innovative work, likely with some of the specific projects presented in Sections 2-5. Because these tasks would be collaborations between scientist participants, it is presumptuous to define them in too much detail today, especially if the Institute were to work well as an incubator of new ideas.

7.1.3. Relevance to facilities

It is not appropriate for the projects to provide support for the operations of facilities, or provide core functions such as basic data reduction or visualization. Facilities have their own software development teams who focus on these problems. The key developers from these facility teams need to be in regular communication with Institute personnel, of course.

An appropriate software project in collaboration with a user facility may be the construction of workflows for interpreting new types of experimental data. This might include tests of a new experimental method, analogous to previous efforts with extended X-ray absorption fine structure (EXAFS), or resonant inelastic X-ray scattering (RIXS), for which



Extracting distance lists from PDF data and constructing candidate local structures. (P. Duxbury)

computations are now essential for interpretation. More likely, Institute software projects will be selected to offer scientific workflows that are needed by an existing technique, but are difficult or restricted in their use today, such as an ab-initio molecular dynamics calculation of the effects of temperature on phonon spectra.

Representation on the Institute Board by members from the national user facilities should help ensure that the software projects are relevant to the needs of X-ray and neutron user facilities. This should also facilitate the integration of new computational workflows into the data analysis methods developed at the national user facilities.

7.1.4. Sustainable Software

The Institute will integrate software packages into workflows, motivated by the scientific opportunities such as described in other sections of this report. Although workflow development will be motivated by science, essential issues for software engineering design were presented in sections 6.1-6.3. It is imperative that many of the workflows developed over the course of scientific collaborations be archived, maintained, and reused in new but related areas of science. Proper attention to design and construction will allow the software to live beyond the initial phases of the Institute itself.

It is expected that the first version of a software workflow will be developed to solve a specific scientific problem that results in a publication. The choice of problems can be made with an eye towards broader scientific applications, so some of the issues of reuse and maintenance can be considered during prototype development. The success of a workflow and the success of its science can then be used to identify those workflows that deserve more thorough engineering.

One of the challenges for developing a software workflow, rather than a standalone software package, is the long-term robustness of their component integration. Best practices will require ongoing evaluation, but today we expect that integration will involve a very-high-level-language such as Python. A runtime-configurable framework may allow for easier maintenance, assuming the inheri-

tance of tools and services can follow the evolution of the tools and the evolution of the component packages. Perhaps some standard design rules applied to user interfaces can facilitate their maintenance, but this is an even greater challenge.

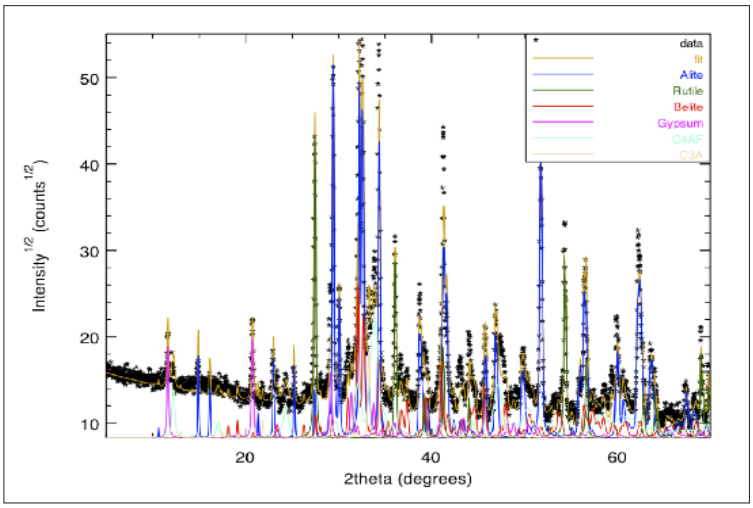
7.2. Governance

Two phases are envisaged for the governance of an Institute. The first is being organized now under the S2I2 conceptualization phase of SIXNS. The P.I. (Fultz) and Co-Is (Billinge, Owhadi, Rehr, Stalzer) are taking the lead to organize the scattering community and to write a proposal that will specify activities for the first three years of an Institute. As part of organizing the scattering community, two additional teams are being formed:

- The first is an Institute Board with executive powers. Membership of the Board will include representatives from the national labs and national user facilities, and an approximately equal number from universities. Members of the Board will help write the initial proposal to the NSF, which will specify the software projects for the first three years of the Institute. During the first budget period of the Institute, the Board will appoint an executive officer who will become the PI for the renewal proposal. The terms of Board members will be planned for continuity, but it is expected that the terms will be approximately four years. A process for electing new Board members needs to be developed during the first years of the Institute.

- The second is a Science Advisory Board (SAB). It will meet approximately annually, perhaps before a national meeting for neutron or X-ray scattering, in years 2, 3, 4 of Institute operations. The members of the SAB will be chosen for their experience in scattering science and/or computational science. International representation is desirable. It is expected that these members will be identified before the first Institute proposal is submitted to the NSF.

The second phase of the Institute would follow a mid-term review of the program by the NSF. At this juncture, Fultz will defer to the Board for selection of the PI for the renewal proposal, and the requirements for redirecting funding to make the renewal viable. It is anticipated that directions will be matched to the skills of the continuing staff members, but there could be a transition to a new set of postdoctoral fellows and graduate students at this time.



Testing the Full Pattern Search-Match method using the Crystallography Open Database (COD). (L. Lutterotti)

The community presently organized for the SIXNS proposal could be important even if the NSF does not fund the Institute through the S2I2 program. Other calls for software infrastructure proposals may likely come from the NSF over the next years. More likely, perhaps, is that the Scientific User Facilities Division of DOE BES will recognize the need for software centers for scattering science. The scale of the proposed staffing (Section 7.6) may prove appropriate for a software development group at a national laboratory. With coordination, several such groups would be able to cover much more of the field of computational scattering science. A national effort to coordinate these activities could emerge from the Science Advisory Board of the SIXNS Institute, for example.

7.3. Institute Activities

A software Institute would bring together leaders in scattering science with leaders in computational science. This is not typical today, and for the emerging field of uncertainty quantification, this would be a first. Activities would be organized around scientific collaborations, where computational and experimental work could have a larger impact than either alone. If based at a university, the individual collaborations would have the style of university research, where a senior investigator would work with postdoctoral fellows and graduate students towards publishable results. A university style of work would need adaptation for building computational workflows that can be reused for future research, accommodating the software development style described in Section 6.1.

An obvious reason for a scattering scientist to collaborate with a computational scattering scientist at the Institute would be to help in software workflow development, documentation and curation of these work products. Nevertheless, initiating a collaboration usually involves surmounting an activation barrier. A member of the Institute staff serving as a "catalyst" would work with the user community to help identify promising projects that could use the existing workflows with appropriate adaptation. Infrastructural support for the collaboration would include brokerage of computational resources, and the maintenance of a software repository and documentation.

An Institute with a total of perhaps 20 professionals, academics and students cannot possibly satisfy the needs of the 14,000 unique users of neutron and X-ray scattering facilities in the U.S. The reason is not so much the numbers of users, but the diversity of their science. With a focus on the physical sciences as proposed in this report, the Institute could provide analysis tools to assist hundreds of users. Other institutes, perhaps affiliated with national user facilities, would be needed to cover more fields of science. Perhaps five to ten such institutes would serve the national need. Nucleating all these at one time seems unlikely today, so starting with one of them is important to do soon. SIXNS could provide experience and models for future institutes. Engaging the user community in the selection of projects will begin with members of the Institute Board, which will include members of the community most involved in scattering research and computation. This group must be expanded to benefit more of the scattering community, giving sufficient attention to individual projects while serving as many researchers as possible. Nevertheless, there will be some time when only a small fraction of the scattering community is served by Institute activities. The Institute needs to work to ensure that broader benefits will emerge over time, and manage community expectations of broader benefit. A steady series of workshops will be part of the Institute operations. Discussion forums are now accepted as a mechanism of user support, and these will be offered and moderated by the Institute.

7.4. Education and Outreach

An institute like SIXNS requires a substantial and effective educational component for success. New users can learn from previous projects as they consider adapting workflows to their own needs, especially if documentation and forums are available. Many workflows could serve as learning tools for young scientists and graduate students.

For graduate students and postdocs, but also for early career professionals such as software developers, exposure to both leading experimental and computational techniques, mentored by leaders in their respective fields, will offer valuable learning opportunities. Institute projects would offer exposure to large-scale experimental facilities, and this has always been useful for graduate students and postdoctoral fellows. Scientists participating in Institute projects will have the additional opportunity to participate in state-of-the-art software development techniques, including object-oriented software design and software development processes.

An Institute has the opportunity to develop on-line resources appropriate for training and education. An effort is emerging in Europe to do this, with emphasis on the quality rather than the quantity of educational material [7.1]. It may be appropriate for the Institute efforts to be merged with this activity to accelerate its development. An early effort, compatible with the European project, may be an introduction to lattice dynamics. A draft of an upper-division graduate-level text has emerged

from the DANSE project [7.2]. Building on textbook coverage of phonons, new content would include the methods of inelastic neutron and X-ray scattering, with reference to and examples from state-of-the-art instruments.

Today, educational content involving simulation software is difficult to fit into existing curricula at academic institutions. It is possible to initiate change by offering resources for homework problems. These need not be focused only on scattering for its own sake, but also on scientific principles in fields of science that benefit from scattering studies. Preparing on-line resources for academic curriculum offers several benefits. It would make the Institute better known in the wider community, and make it easier to include scattering topics in university curricula. Easy access to quality training resources will broaden the user base of national large-scale facilities. Furthermore, the students and postdocs involved in preparing the material will gain valuable experience for careers in academia.

With the changing demographics of the U.S. population, improving diversity within the scattering community and the Institute can ensure that scattering science remains vigorous in the future. An Institute could participate in an outreach effort such as the LEAD (Learning, Education and Development) program at Caltech, which exposes exceptional underrepresented minority high school students to the excitement and rigor of engineering research, especially computing, through a three-week summer residence on the Caltech campus [7.3]. The program is supported in part by Intel and Google, but an Institute could contribute to its scope and breadth.

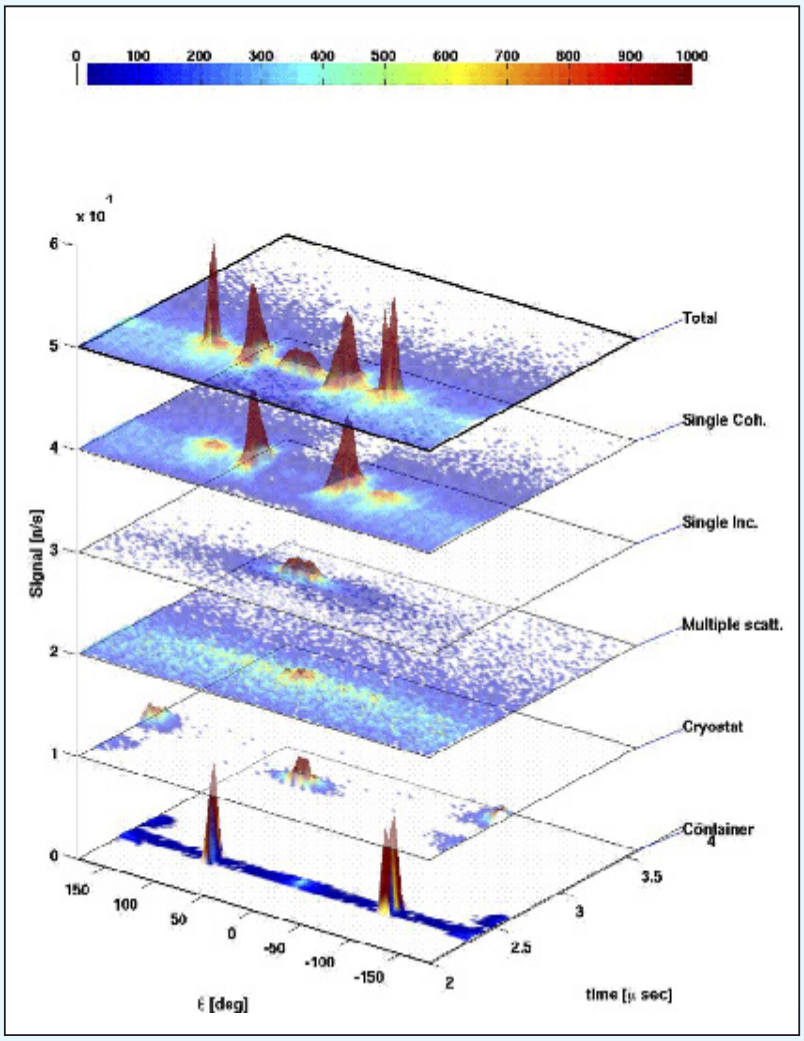
An important problem in computational scattering science is the low representation of women in both computer science and in physical science. The investigators have made efforts to address this issue at the graduate student level, and the Institute would make available opportunities at the postdoctoral level. The subsequent career stages need attention, especially for permanent staff positions at universities or national laboratories. First, a career path in computational scattering science is unusual. Second, there are few women or underrepresented minorities who have taken this path to a permanent position. These twin issues are discussed next.

7.5. Career Paths

SIXNS would benefit the careers of students, postdocs and early-career professionals by allowing them to work with state-of-the-art scattering methods and modern methods of computational materials science. Their mentoring would be by world-leading experts, and there would be networking opportunities in academia and national and international research facilities. At the postdoctoral level, many experimental research groups value individuals who can use the tools of modern computational materials science. The SIXNS Institute can offer excellent opportunities for junior computational scientists to do new science by collaborating with experimental groups. There will be a substantial expectation for these junior scientists to publish new science to advance their careers, and through collaborative work with scattering scientists this expectation can be fulfilled. Most long-term career paths are through the established academic fields of science like chemistry, physics, and materials science. Computational science is setting new directions in these fields. Proven skill in software and experiment design, publication of scientific results, documentation, distribution and maintenance of codes should be valuable for academic, research, or industrial career paths outside the Institute. From our experience with other software projects, this promise is only partly realized for the following reasons.

University hiring in departments of physics, applied physics, chemistry and materials science tends to focus on phenomena in materials, matter, and energy. Faculty positions are viable career paths for postdocs with a strong focus on the materials research aspects of their work. Many postdoctoral fellows in X-ray and neutron scattering research at national laboratories have moved to university departments in recent years. The computational scattering science aspect of the work should help, provided the main focus is still on the science and not on the software. Good mentoring is crucial if the postdoctoral fellows in the Institute are to plan research careers strategically.

Other young scientists in the Institute may gravitate towards software and computing, which offer their own challenges and rewards. A strong emphasis on modern software development is



Virtual experiments in neutron scattering -- simulated data from liquid Ge virtual experiment. (L. Udby)

not a tradition for physical scientists, as opposed to electronic design or cryogenic expertise, for example. A position of "computational scattering scientist" may be appropriate at a national user facility, parallel to the position of instrument scientist, or a scientist responsible for sample environment or detector development. Although the position of computational scattering scientist is a new concept, such persons could serve the strategic need of a national laboratory by helping to connect its high-performance computing activity to its national user facility.

Some young persons working on computational science will migrate away from physical science entirely. Career paths for theorists have always been challenging, and many have found careers in other sectors of the economy. Sometimes the opportunity to change path towards software design and high-performance computing is appealing to young persons, and an Institute must provide mentorship for young persons interested in this path. Although other fields may prove more lucrative, they may return to scientific computing later.

7.6. Staffing and Budget

The scope of activities is bounded by budget. Most of the budget for a software Institute would be salaries. For a nominal budget of 3 M$ per year, support could be:

- Some salary offset for PI, Co-I (as required by law)

- (1) Software architect
- (1) Computer scientist
- (3) Developers/Scientists
- (1) Catalyst
- (7) Postdocs
- (5) Graduate students
- (1) Sysadmin
- (1) Education/outreach specialist
- (2) Administrators (local, external focus)
- (1) cluster of order 1,000 cores
- (2) smaller systems for architecture tests
- Licenses, hardware for builds, repository, office
- Participant support
- Travel

7.7. Metrics for Evaluation

For evaluating the success of software projects, there is widespread concern that the usual metrics of publications, citations, and downloads are only partly successful. For example, major parts of a code base can be merged into the development trunk of other projects with little recognition, or key ideas from prototypes can be used in fresh efforts that have high impact. In part for this reason, we propose that the metrics for the software Institute be primarily those of scientific accomplishment, for which the metrics are better established. Adherence to these conventional metrics is probably better for the careers of most of the postdoctoral fellows and graduate students who participate in the work, at least those who pursue a career in science.

An emphasis on science will also motivate others in the X-ray and neutron scattering community to take interest in the Institute when the work of the Institute produces scientific results. A good feature of the scattering science community is that it already has an existing culture that accommodates collaborations between scattering scientists and other scientists. This extends naturally to members of the Institute, including those early in their careers. We expect even more collaborations as useful workflows and expertise are accumulated by the institute, so the institute should support a wider range of activities than at its inception. The rates of publications, citations, downloads, and size of the user base will grow, giving a positive trend over time.

There has been much community lamentation about the use and abuse of the H-index and the publication of results in high-impact journals. For better or worse, however, these simple criteria will remain for the next few years, and the Institute must allow young scientists to optimize their scores by these metrics. Focusing the work of the young computational scientists on collaborative projects that elevate the quality and impact of scattering measurements should be a way to do so.

[7.1] http://prezi.com/nbw6wr0-ytzc/vnt_garching2012/?utm_source=website&utm_medium=prezi_landing_related&utm_campaign=prezi_landing_related_author

[7.2] http://www.cacr.caltech.edu/projects/danse/doc/Inelastic_Book.pdf

[7.3] http://www.leadprogram.org/summer-programs/lead-engineering-sei-csi/caltech

Participants in the Workflows for Computational Scattering Science 2013

Michael Aivazis
California Institute of Technology

Ercan Alp
Argonne National Laboratory

Jim Beck
California Institute of Technology

Simon Billinge
Columbia and Brookhaven

Katheryn Broersma
California Institute of Technology

Daniel Brooks
California Institute of Technology

Sharon Brunett
California Institute of Technology

Yong Cai
Brookhaven National Lab

Roberto Car
Princeton University

Alessandro Cunsolo
Brookhaven National Laboratory

Pengcheng Dai
The University of Tennessee

Paul Dawson
Cornell University

Ewa Deelman
USC Information Sciences Institute

Olivier Delaire
Oak Ridge National Laboratory

Przemyslaw Dera
University of Chicago

Tom Devereaux
SLAC National Accelerator Laboratory
and Stanford University

George Djorgovsky
California Institute of Technology

Phillip Duxbury
Michigan State University

Emmanuel Farhi
Institut Lague-Langevin

Brent Fultz
California Institute of Technology

Roger Ghanem
University of Southern California

Matthew Graham
California Institute of Technology

Mark Green
Tech-X Corporation, Janus Systems Corporation

Mark Hagen
SNS/ORNL

Michael Hu
Argonne National Laboratory

Mark Hybertsen
Brookhaven National Laboratory

Kevin Jorissen
University of Washington

Daniel Katz
National Science Foundation

Dennis Kim
California Institute of Technology

Jiao Lin
California Institute of Technology

Luca Lutterotti
University of Trento

Anders Markvardsen
ISIS, STFC, UK

Noa Marom
Tulane University

Lisa Mauger
California Institute of Technology

Michael McKerns
California Institute of Technology

Paul Messina
Argonne National Laboratory

Jorge Muñoz
Intel Corporation

Reinhard Neder
University Erlangen

Matthew Newville
University of Chicago

Raymond Osborn
Argonne National Laboratory

Houman Owhadi
California Institute of Technology

Lars Pettersson
Stockholm University

David Prendergast
Lawrence Berkeley National Laboratory

Thomas Proffen
ORNL

John Rehr
University of Washington

Dmitry Reznik
University of Colorado-Boulder

Daniel Roach
University of Salford

Anthony Rollett
Carnegie Mellon

Stephen Rosenkranz
Argonne National Laboratory

Galen Shipman
ORNL

Hillary Smith
California Institute of Technology

Mark Stalzer
California Institute of Technology

Julian Tao
China Spallation Neutron Source

Brian Toby
Argonne National Laboratory

John Tranquada
Brookhaven National Laboratory

Matt Tucker
ISIS, STFC, UK

Linda Udby
University of Copenhagen

Nathalie Vast
Commissariat a l'Energie Atomique

Sven Vogel
LANL

Nancy Wilkins-Diehr
San Diego Supercomputer Center

Andrew Wills
UCL

Taner Yildirim
NIST and UPENN

Cai Yong
Brookhaven National Laboratory

Linda Young
Argonne National Laboratory

Colophon

This document was assembled using Adobe InDesign.
Frutiger, Myriad Pro and Adobe Garamond Pro are the three font families used.

•

Design and layout for this report was done at Caltech Graphic Resources in November 2013.